An Advanced A-V Player to Support Scalable Personalised Interaction with Multi-Stream Video Content

Zhenchen Wang

Submitted in partial fulfilment of the

requirements for the degree of

Doctor of Philosophy

London, June 2011

Abstract

Current Audio-Video (A-V) players are limited to pausing, resuming, selecting and viewing a single video stream of a live broadcast event that is orchestrated by a professional director. The main objective of this research is to investigate how to create a new custom-built interactive A-V player that enables viewers to personalise their own orchestrated views of live events from multiple simultaneous camera streams, via interacting with tracked moving objects, being able to zoom in and out of targeted objects, and being able to switch views based upon detected incidents in specific camera views. This involves research and development of a personalisation framework to create and maintain user profiles that are acquired implicitly and explicitly and modelling how this framework supports an evaluation of the effectiveness and usability of personalisation.

Personalisation is considered from both an application oriented and a quality supervision oriented perspective within the proposed framework. Personalisation models can be individually or collaboratively linked with specific personalisation usage scenarios. The quality of different personalised interaction in terms of explicit evaluative metrics such as scalability and consistency can be monitored and measured using specific evaluation mechanisms.

Acknowledgements

I would like to thank following people:

My supervisor Stefan Poslad, for his encouragement and confidence in my work. For advice, moral and practical support when I needed it. Stefan's support made it possible for me to complete this work.

My second supervisors, Alan Pearmain and Nick Bryan-Kinns, who were involved in my supervision, for helpful discussions and research ideas.

BBC RandD colleagues, Peter Brightwell, Hannah Fraser, Maxine Glancy, for making video content available during the time my research was carried out and for making valuable suggestions and ideas on usability testing.

Atos Origin RandD, INOV and other colleagues of the EU FP7 My eDirector 2012 Project including Elena Garrido, Nelson Escravana who enlarged my research vision to an industrial level.

All those who offered me opportunities to extend my scientific knowledge during last four years, these include Nigel Chappell, Shihua Wang, Dejian Meng, Kraisak Kesorn, Zekeng Liang, Haibo Mei, Uko Asangansi.

My parents Yunqing Wang and Ailan Liu, for their love and support.

The research leading to these results has received funding from the European Union's Seventh Framework Programme ([FP7/2007-2013]) under grant agreement No. ICT-215248 and from Queen Mary University of London.

List of Abbreviations

A-V	Audio and Video
BMP	Bayesian Point Machine
CDN	Content Distribution Network
CTTE	ConcurTaskTrees Environment
DT	Decision Tree
DVR	Digital Video Recorder
FCM	Fuzzy C-means
GOMS	Goals, Operators, Methods and Selection rules
HCI	Human–computer interaction
HMM	Hidden Markov Model
HTA	Hierarchical task analysis
ICT	Information and Communication Technology
IHCI	Implicit HCI
IIS	Internet Information Server
MD	Motion Difference
OH	Objects Highlighting
PP	Prediction Precision
RIA	Rich Internet Application
RP	Replay
SI	System Interface
STB	Set Top Boxes
UI	User Interface
WSR	Wilcoxon signed-rank test
ZUI	Zoomable User Interface

Table of Contents

1 Introduction 1
1.1 Motivation 1
1.2 Thesis Statement 3
1.3 Research Objectives 4
1.3.1 Research Objectives and My eDirector 2012 Project 4
1.4 Research Novelty 5
1.5 Thesis Outline 5
2 Literature Review 7
2.1 Audio -Video Players 7
2.2 User Modelling including Interaction and Task Modelling 10
2.2.1 Cognitive and Empirical User Models 11
2.2.2 User Task Models 12
2.2.3 Comparison of User Models 14
2.3 Personalised Viewing of Sports Events 15
2.3.1 Content Acquisition and Production 15
2.3.2 User Model Acquisition: Preferences 18
2.3.3 User Acquisition Model, Interactions and Tasks 19
2.3.4 Adapting User Content to User Models 27
2.4 Personalisation Evaluation 31
2.4.1 Comparison based Approach 32
2.4.2 Knowledge based Approach 34
2.4.3 Hypothesis Testing Based Approach 35
2.5 Summary 37
3 User Interaction in a Next Generation A-V Player 39
3.1 Interaction Requirements 39
3.2 Interactive User Task Model 40
3.2.1 Task Description 41
3.2.2 Personalisation Implication from Task Model 45
3.3 Task Model Application 47
3.3.1 Sports Events Selection Task Model 48
3.3.2 Multi-Angle Viewing Task Model 50
3.3.3 Selective Target Zooming Task Model 52

3.3.4 Time-Shift Viewing Task Model 54
3.4 Task Interface Requirements for Supporting Personalisation 56
3.4.1 Interface Requirements in Multi-Angle Viewing Task 56
3.4.2 Interface Requirements in Selective Target Zooming Task 57
3.4.3 Interface Requirements in Time-Shift Viewing Task 57
3.5 Summary 57
4 Personalisation in a Next Generation A-V Player 59
4.1 Personalisation Objectives 59
4.2 Personalisation Framework Overview 60
4.2.1 Personalisation Design Issues 62
4.2.2 User Model or Profile 66
4.3 Personalised Event Selection 68
4.3.1 User Model 71
4.3.2 Sports Events Selection based upon Individual Recommendations 71
4.3.3 Sports Events Selection Based upon Group Recommendations 76
4.4 Personalised Multi-Angle Viewing 85
4.4.1 Multi-Angle Viewing User Model 85
4.4.2 Multi-Angle Viewing Personalisation Model 86
4.5 Personalised Selective Target Zooming 97
4.5.1 Selective Target Zooming User Model 97
4.5.2 Selective Target Zooming Personalisation Model 98
4.6 Time-Shift Viewing 107
4.6.1 Time-Shift Viewing User Model 107
4.6.2 Time-Shift Viewing System Model 107
4.7 Summary 114
5 Evaluating Personalisation Performance in a Next Generation A-V Player 115
5.1 Overview 115
5.2 Test Parameter Acquisition 116
5.3 Personalisation Evaluation at Run-Time 117
5.4 Consistency Test 118
5.5 Scalability Test 120
5.6 Passive and Active Personalised Interaction Evaluation 121
5.6.1 Passive System Adaptation 121
5.6.2 Active System Adaptation 122
5.7 Summary 123

6 Next Generation A-V Player Implementation and Evaluation 124
6.1 Personalisation Implementation 124
6.1.1 Overview 124
6.1.2 Personalised Sports Events Selection (Individual Recommendation) 124
6.1.3 Personalised Sports Event Selection (Group Recommendations) 125
6.1.4 Personalised Multi-Angle Viewing 125
6.1.5 Personalised Selective Target Zooming 127
6.1.6 Time-Shift Viewing 129
6.2 Personalisation Evaluation 130
6.2.1 Personalised Sports Events Selection (Individual Recommendations) 130
6.2.2 Personalised Sports Event Selection (Group Recommendations) 141
6.2.3 Personalised Multi-Angle Viewing 146
6.2.4 Personalised Selective Target Zooming 154
6.2.5 Time-Shift Viewing 166
6.3 Usability Evaluation 172
6.4 Summary 178
7 Achievement and Future Work 179
7.1 Achievements 179
7.1.1 Development Achievements 180
7.1.2 Personalisation Achievements 180
7.1.3 Publications 181
7.2 Further Work 182
Appendix A 184
Appendix B 187
Appendix C 190
Bibliography 192

Table of Figures

Figure 3-1 Graphical representation of a sports events selection task model 43
Figure 3-2 Event selection task in Kripke structure 44
Figure 3-3 Task model annotated with CTL transition path possibility 46
Figure 3-4 Task nodes with information entropy values for a 'Browsing' task that requires personalisation 47
Figure 4-1 Personalisation in interactive systems 59
Figure 4-2 Personalisation framework 62
Figure 4-3 Use model in a personalised interactive system 67
Figure 4-4 Sports events selection personalisation model 69
Figure 4-5 Personalised event selection user model 71
Figure 4-6 Traditional individual user feedback driven personalisation model 72
Figure 4-7 Events Importance Continuum 73
Figure 4-8 Group driven personalisation model 78
Figure 4-9 Bayesian network 84
Figure 4-10 Personalised multi-view angle viewing user model 86
Figure 4-11 System architecture for a personalised multi-angle viewing system 87
Figure 4-12 Personalised multi-angle viewing model 88
Figure 4-13 Camera switching personalisation model 93
Figure 4-14 Switching intervals fuzzy sets 94
Figure 4-15 Camera switching HMM topology camera types 96
Figure 4-16 Selective target zooming task user model 97
Figure 4-17 System architecture for a personalised ZUI 99
Figure 4-18 Selective target zooming personalisation model 100
Figure 4-19 Personalised Zooming Control Model 103
Figure 4-20 Time-shift viewing user model 107
Figure 4-21 Time-shift viewing system model 108
Figure 4-22 Motion change distribution with respect to a Gaussian distribution 110
Figure 4-23 Motion change matrix 111
Figure 4-24 Highlight incidents model 112
Figure 4-25 Object highlighting model 113
Figure 5-1 Personalisation evaluation model 116
Figure 5-2 Personalisation quality aspects and sources of evaluation parameters 117
Figure 5-3 Runtime personalisation evaluation workflow 118
Figure 5-4 Personalisation consistency evaluation algorithm workflow 119
Figure 5-5 Personalisation scalability evaluation algorithm workflow 120
Figure 6-1 Events browsing UI 125
Figure 6-2 Multi-angle viewing UI – camera selection 126
Figure 6-3 Multi-angle viewing UI - camera switching 126
Figure 6-4 Zooming animation implementation 127
Figure 6-5 Video ZUI 128
Figure 6-6 Decision rule of system actions according to the detected motion conditions 130
Figure 6-7 Object visual annotation UI 130

Figure 6-8 Supporting camera settings 147
Figure 6-9 Bitrates changes without multi-stream adaptation 149
Figure 6-10 Bitrates changes with multi-stream adaptation 149
Figure 6-11 Playback bitrates change standard deviation comparison 150
Figure 6-12 Time latency between zooming events and bitrates changes 157
Figure 6-13 A Comparison of visual quality between non-bitrate adaptation and bitrate adaptation 158
Figure 6-14 Time latency for time-shift playback 159
Figure 6-15 400m Event objects highlighting histograms vs. visual annotation 168
Figure 6-16 Long jump event objects highlighting histograms versus visual annotation 169
Figure 6-17 Incidents and highlights are plotted onto the event timeline (time unit = second) 169
Figure 6-18 Start running replay scene (first frame) 169
Figure 6-19 Third corner turning replay scene 170
Figure 6-20 Last 100m close up replay scene 170
Figure 6-21 First Athlete crosses the finish line replay scene (last frame) 170
Figure 6-22 First OH/RP point 171
Figure 6-23 Second OH/RP point 171
Figure 6-24 Third OH/RP point 171
Figure 6-25 Fourth OH/RP point 172
Figure 6-26 Player main page 173
Figure 6-27 Questionnaire page 173
Figure 6-28 PCA factor loadings 176
Figure 6-29 PCA Factor scores for each user 176
Figure 6-30 PCA Factor scores for users grouped w.r.t. Internet TV use frequency. 177
Figure 6-31 PCA Factor scores for users grouped w.r.t. age 177
Figure 6-32 PCA Factor Scores for each users grouped w.r.t. gender 177

Table of Tables

Table 2-1 User Tasks, Personalisation in A-V Players 9
Table 2-2 Pros and cons of ZUI techniques 22
Table 2-3 Pros and cons of camera switching approaches 25
Table 2-4 Comparison based evaluation designs 33
Table 2-5 Knowledge based evaluation designs 34
Table 2-6 Evaluation approach and solutions to challenges 36
Table 2-7 Hypothesis testing steps 36
Table 3-1 Task Models Evaluation Matrix 41
Table 3-2 Task notations and relations notations 42
Table 3-3 Task Relations as CTL Expressions 45
Table 3-4 Algorithm to identify critical sub task 46
Table 3-5 Sports events selection task model 49
Table 3-6 Multi-angle viewing task model 51
Table 3-7 Selective target zooming task model 53
Table 3-8 Time-shift viewing task model 55
Table 3-9 Critical interactive task for personalisation and associated task interface requirements 58
Table 4-1 Design choices 63
Table 4-2 Sports Events Stages 73
Table 4-3 Football Event Metadata Description 74
Table 4-4 Metadata Description of other Events 75
Table 4-5 User Information with predefined values 79
Table 4-6 Sports attributes and values 80
Table 4-7 Screen resolution bitrate adaptation algorithm 90
Table 4-8 Multi-stream adaptation algorithm 91
Table 4-9 Camera Switching User Profile Encoded in XML 93
Table 4-10 Viterbi algorithm 97
Table 4-11 Video quality adaptation algorithm 101
Table 4-12 Personalised zooming user profile in XML 102
Table 4-13 Personalised zooming control cluster algorithm 105
Table 4-14 FCM clustering algorithm 106
Table 4-15 Future zooming region of interest determination algorithm 106
Table 4-16 Interaction and live editing task interface 109
Table 6-1 Sports events used in evaluation experiments 131
Table 6-2 Simulated live schedule 131
Table 6-3 12 Users' recommendation accuracy values 132
Table 6-4 Recommendation accuracy values for operational evaluation, shaded values are pseudo RA values for statistic calculation purpose 133
Table 6-5 Hypothesis testing with fixed RA median of 0.6 operational, p=0.05 134
Table 6-6 Hypothesis testing with fixed RA median of 0.7 operational, p=0.05 135
Table 6-7 Hypothesis testing with fixed RA median of 0.8 operational, p=0.05 136
Table 6-8 Hypothesis testing with previous RA median, p=0.05 137

Table 6-9: Apply rule 6.2.1.2-1 to user 12 to decide to update the current recommendation accuracy threshold value 138
Table 6-10 Hypothesis testing with previous recommendation accuracy produced by random approach, p=0.05 139
Table 6-11: Apply Rule 6.2.1.2-2 to user 5 to decide to activate the personalisation 140
Table 6-12 Correlation coefficient for number of uses and recommendation accuracy. 141
Table 6-13 Recommendation accuracy results for schedule session 2 142
Table 6-14 Recommendation accuracy results for schedule session 3 143
Table 6-15 Recommendation accuracy results for schedule session 4 143
Table 6-16 Hypothesis testing for a single personalization session with 10 X resampling (Schedule session 2) 144
Table 6-17 Hypothesis Testing for a Single Personalization Session with 10 X resampling (Schedule session 3) 144
Table 6-18 Hypothesis testing for a single personalization session with 10 X resampling (Schedule session 4) 145
Table 6-19 Mean recommendation accuracy with corresponding user number 146
Table 6-20 Retrieved PP for a camera switching system with personalisation versus a generated random PP on system without personalisation 147
Table 6-21 Hypothesis testing with PP median, p=0.05 150
Table 6-22 Applying rule 6.2.3.3-1 to user 5 to decide the PP threshold value 151
Table 6-23 Hypothesis testing with PP median produced by random approach, p=0.05152
Table 6-24 Applying Rule 6.2.3.3-2 to decide to activated auto camera switching for users 4 and 5 in Table 6-21 152
Table 6-25 Correlation coefficient of number of uses and prediction precision 153
Table 6-26 Accumulated prediction precision values of 6 users in 11 personalised zooming sessions for the long jump event 155
Table 6-27 Accumulated prediction precision values of 6 users in 11 non-personalised zooming sessions for the long jump event 155
Table 6-28 Accumulated prediction precision values of 6 users in 11 personalised zooming sessions for the 400m event 156
Table 6-29 Accumulated prediction precision values of 6 users in 11 non-personalised zooming sessions for the 400m event 156
Table 6-30 Hypothesis testing with a previously accumulated PP median, for the long jump event, p=0.05 160
Table 6-31 Hypothesis testing with previous accumulated PP median, p=0.05 (400m) 161
Table 6-32 Applying Rule 6.2.4.3-1 to decide the threshold PP value on user 2, 3, 4 and 5 (Table 6-30) 162
Table 6-33 Applying Rule 6.2.4.3-1 to decide the threshold PP value on user 1, 4 and 5(Table 6-31) 162
Table 6-34 Hypothesis testing with a previously accumulated PP median produced by random approach, for the Long jump event p=0.05 163
Table 6-35 Hypothesis Testing with a previous accumulated PP median produced by a system without personalisation for the 400M event, p=0.05 164
Table 6-36 Applying rule 6.2.4.3-2 to decide activation of personalised zooming on User 1 (Table 6-35) 165
Table 6-37 Correlation coefficient for a number of uses and prediction precision 166

Table 6-38 Incidents highlighting test results ... 167
Table 6-39 Incident highlighting precision in terms of lead time ... 168
Table 7-1 Player features comparisons ... 180
Table 7-2 Personalisation implementation status and evaluation results ... 181

Chapter 1

1 Introduction

Existing audio and video players such as customised players in set-top boxes and soft players such as the windows media player support relatively little interaction to manipulate live video content. Although Rich Internet multimedia Applications (RIA) can enrich the user interaction, the user interaction and ability to personalise the interaction is still limited. Next generation broadcast sports video content services offers a much greater potential for user interaction and personalisation, e.g. the My-e-Director 2012 EU FP7 project (http://www.myedirector2012.eu) enables viewers to view sports events from multiple camera angles, to select the camera angle, to selectively track moving objects in the video stream and perhaps to switch cameras based upon detected sports incidents. In this thesis, and within the My eDirector 2012 project, the focus is on personalising much richer A-V player interaction. This is akin to viewers being able to interactively direct the views of an event. New interaction options could include user-centred camera switching, zooming on targets of interest, slow motion control etc. However, personalisation of these interactions can be difficult. Traditional user task modelling approaches cannot adequately describe personalised interactive user tasks because they tend not to define an explicit model to track user interaction and how to adapt tasks to this interaction. (See Section 2.1) Even when the personalisation is taken into account in task modelling, additional challenges can be introduced to achieve such personalisation. (See Section 2.3) In addition, at present there is no suitable evaluation framework for adaptive personalisation (See Section 2.4), which often makes the evaluation of personalisation a challenge.

1.1 Motivation

A main incentive to view live sports events online is to enhance the viewing experience. Conventional TVs and Set Top Boxes (STB) contain inbuilt media controllers that support proprietary features; are not easy to iteratively upgrade; that conventionally use a remote controller for interaction, and that offer limited multi-view support. In contrast, online Web-based A-V players can offer users richer interaction options to manipulate the video content playback via on-screen widgets and multimode input support, easing content navigation, e.g. mouse, keyboard, touch, remote controllers, and such players can be more easily upgraded.

This also raises new challenges, such as which types of user interaction to support, the availability of many more channels, more flexible simultaneous channel viewing and the ability to create mashups with other media related to the video channels. When a system needs to adapt to user interaction, it may also be important for a system to know where it should assist a user during an interactive task. For example, if a user needs to perform the following steps or actions 'open menu', 'select sports event items' and 'confirm selection' to complete an event selection task, then the system needs to determine in which of the steps it should help a user. In addition to this, a user model or profile that describes user preferences is required in order to allow the system to do the adaptation. To be useful, user profiles may need to be maintained in order to be dynamic to adapt to specific viewing contexts and to changing user preferences.

Online viewing systems can inherently provide access to many more sports events at any given time. Multiple concurrent sports event channels may be on offer, multi-camera views of a particular sports event can be streamed and higher quality in addition to professional content can be streamed. Traditionally, one single video stream is displayed. Multiple video streams can also be displayed, offering richer video content. However, multi-stream screens can be distracting or may reduce the resolution as the screen real-estate is shared between multiple views and multiple screens may consume more network bandwidth.

Viewers may find it difficult to navigate through menus that pertain to a large content selection. Reducing the selection though filtering the channels on offer to only match user preferences to the live sports events in progress is a common design to handle the issue of too much choice. However, matching a user profile to a channel tends to be very coarse grained. Note also the difference between a channel and an event here. A channel may schedule a time sequence of a range of different types of sports event, commentaries, interviews and news items – many parts of a channel schedule may not match a user profile. In addition, viewers' preferences of sports event change as events progress, e.g., more viewers watch the progress of winners versus losers and finals rather than preliminary rounds. This makes it problematic to use some high-level video channel content metadata, such as 'competition round' to label content and to match it to user preferences. Hence, a new user-centred recommender is required. Such a recommender should allow the system to maintain users' changing preferences and to match them to many instances of video content. The design of such a recommender should also balance control of the preference-to-content matching through allowing a degree of user control to determine when recommendations are needed.

Traditional broadcast channels that schedule content in some state of the art live sports event viewing systems are mainly determined and composed by a human director. For live events, multiple video streams are available in the human director's control room and identified video objects are filtered and visualised by professional users led by a human director. This human director driven approach may not be a scalable solution when many events occur concurrently when there are not many directors. An alternative solution is to allow users to have full control of directing the schedule of content of their own events. However, this introduces additional issues such as being able to repeat and complete operations consistently, e.g. highlight sports incident and having expertise to direct the content with these operations. Therefore, in order to allow users to direct their own events more easily while still viewing professionally directed viewing content, a directing supporting system is required.

Evaluating the effectiveness of personalising live events is challenging. The conventional way to evaluate personalisation is to undertake a summative evaluation of a final system and to directly seek feedback from users via a usability questionnaire. This type of evaluation is of limited use to evaluate context-driven, e.g., that depends on time-dependent sports content, adaptive system behaviour. During live sports events, personalisation may be applied in short in order to dynamically adapt to user's changing preference while events progress. As a result, a supplementary evaluation model that is used in situ during viewing and takes into account time-dependent implicit user feedback is also needed to effectively evaluate personalisation.

1.2 Thesis Statement

Existing A-V players have limitations in offering advanced user interaction to view live events such as sports events. Moreover, they do not provide personalised services that allow a player to adapt to users' preferences. This thesis researches and develops a next generation A-V player that will offer more advanced user and system interaction and dynamic personalisation.

1.3 Research Objectives

The main research question this thesis seeks to address can be summarised as how to enrich the interaction of an advanced live sports audio-video player and how to personalise this in a usable way? This research question also concerns the issue of how to allow an A-V player to alter its functionality to adapt to users' preferences and how to allow the system to better manage the performance of the delivered personalised A-V services. In order to address the issues discussed in the motivation section, the research objectives can be summarised as follows:

- Research and develop the use of personalisation models to enrich the interaction with a live sports audio-video player sub-system. This objective can be further broken down into:
    - Research and development of a task model that supports personalisation (Chapter 3)
    - Research and development of a user model that allows the system to adapt to dynamic user preferences (Chapter 4)
    - Research and development of a personalisation mechanism to facilitate richer interaction with live sports content within specified user interaction tasks (Chapter 4).
- Research and develop the use of an evaluation model for an interactive live sports viewing system that is able to determine the effectiveness of the interaction and personalisation with respect to explicit notions of specific human-computer interaction (HCI) criteria (Chapter 5 & Chapter 6).

1.3.1 Research Objectives and My eDirector 2012 Project

This first objective represents the author's contribution to the My eDirector 2012 project that partially funded his research. The second objective arises from the first objective, however this objective, validating personalising interaction, was not a specific focus of the My eDirector 2012 project but represents a further contribution by the author. The research and development work to achieve these objectives presented in this thesis was solely the work of the author. In this thesis, only part of the general user requirements for the My eDirector project, was taken from a user survey representing work undertaken by the project as a whole (section 3.1), was used for the A-V player, as indicated by a reference.

1.4 Research Novelty

An overview of the novelty of this work is listed below:

- A new user task model has been researched and developed that enables advanced personalised interaction for a next generation A-V player.
- A personalisation model is advanced to enable a customised interactive A-V player to:
  - Model both individual and group users' preferences in the context of live sports events and to support active recommendations to groups with a low group re-clustering overhead.
  - Model and leverage individual user's preferences to actively select targets to zoom in on and to use multi-angle views
  - Automatically adapt views of pre-determined sports incidents, through highlighting scenes and objects.
- Research and develop a system to actively adapt video content delivery to a network context:
  - Bitrate adaptation handles bandwidth contention that occurs when two or more streams are concurrently displayed on a screen during zooming and during multi-camera viewing
  - Time-shift viewing can be interleaved into live events to support replays and objects highlights.
- Define and apply a new evaluation model to assess personalisation performance with respect to consistency and scalability.

1.5 Thesis Outline

The thesis is structured as follows. Chapter 2, Literature Review, reviews the existing work on personalised interaction with the focus on sports events viewing systems. Different aspects of a personalised interactive system are reviewed in terms of interactive task models, personalisation of interactive sports event viewing and the evaluation of such personalisation. The analysis indicates that existing approaches seem to be unable to support a personalised interactive system for live sports event viewing.

Chapter 3, Interaction in a Next Generation A-V Player, addresses fundamental design issues concerning a personalised interactive system, including: the identification of domain based interaction requirements during live sports viewing; the design of a new user task model that supports personalisation and the application of such model in a sports viewing system.

Chapter 4, Personalisation in Next Generation A-V Player, personalises the interactive tasks based upon a proposed personalisation framework. The personalisation framework defines the required personalisation within the context of live sports viewing. In terms of the defined personalisation framework, personalisation models are defined to support; live sports events selection, camera switching within a live sports event, selective target zooming and time-shift (slow motion) viewing. In addition to enhancing personalisation for these defined interactions, the proposed models also address some novel issues beyond the state of the art such as the use of multi video stream bitrate control that could otherwise degrade the use and usability of personalisation in practice.

Chapter 5, Evaluating Personalisation Performance in Next Generation A-V Player, proposes a new evaluation method to effectively evaluate the performance of the personalisation. The proposed evaluation model uses a Hypothesis Testing based approach and is able to evaluate the personalisation at different stages including the operational (i.e. within a system use session) and post-operation stage (i.e. after a system use session). Metrics such as personalisation scalability and personalisation consistency are defined.

Chapter 6, Next Generation A-V Player, Implementation and Evaluation, applies the proposed personalisation evaluation method, defined in chapter 5, within a next generation A-V player. Both the proposed evaluation model and usability are evaluated in into lab trials.

Chapter 7, Achievements and Future Work, summarises the novelty and, achievements of the work reported in this thesis. Further work to improve the work done is also proposed.

Chapter 2

2 Literature Review

The literature review in this chapter covers specific issues related to next generation A-V players. The state of the art of audio/video terminal is firstly reviewed and followed by a review of the interactive task and user models. Then, personalisation is surveyed because personalisation enriches and tailors users' interactions with A-V payers. Personalisation is surveyed in terms of the approach used, user preference elicitation, interactions and user preference adaptation. Finally, the personalisation evaluation approaches are reviewed.

2.1 Audio -Video Players

This work focuses on the research and development of an advanced A-V player for live sports events. In support of the thesis statement (section 1.2), the state of the art of A-V players are firstly surveyed and analysed in terms of their support for both basic and advanced features and personalisation. The features are the basic A-V controls (play, stop, volume control, etc.), advanced A-V controls (Fast Forward or FFW, Slow Motion, Live Pause), video content selection modes (i.e. channel vs. content), and single-screen view vs. multi-screen views. Personalisation support includes a provision for both passive and active adaptation to user's preferences.

NoTube is a European project exploring television's future in the ubiquitous Web (NoTube Consortium, 2009) (Aroyo et al., 2009). NoTube's main contribution to leading edge video player is to create a new generation of Web services for context dependent, personalised, selection and presentation of TV content. This shifts digital entertainment from a single-TV viewer activity to a community-based experience through sharing preferences. It realizes distributed personalisation in an interactive and multi-device environment, enabling anywhere and anytime TV entertainment with the ubiquitous Web. User preferences are mainly determined by the metadata describing the video content such as the content genre. While this approach is often used in video retrieval systems, it can be less effective when used in live video broadcast system as there can be less or no live metadata available and the video streams can be bound to cameras rather than to TV channels.

In the **MYMEDIA** project (MYMEDIA Consortium, 2009), the problem of information overload is investigated. A dynamic personalisation of multimedia content approach is used to solve this problem. The dynamic personalisation concept draws on ideas from some commercial on-line services such as Amazon's recommender service that recommends content from more than one supplier. In the MYMEIDA project, an implicit user model (Rendle et al., 2009) is proposed to retrieve personalised information using a client side application parasitizing on an existing Web searching engine. Unlike traditional recommender systems that use a content catalogue from one service provider, MYMEDIA integrates the recommendations from different content providers. The core component of the system is the content catalogue protocol that enables multiple recommendation systems to be integrated. However, because user preferences are primarily retrieved using search keywords, the matching between user preferences and recommendations from content providers can be inaccurate when there are semantic variations or uncertainties on the keywords.

The **MOBILE3DTV** project (MOBILE3DTV consortium, 2009) aims to develop suitable stereo-video content-creation techniques and includes gathering new knowledge about user experiences in terms of user acceptance and satisfaction with mobile 3DTV content. The user experience in MOBILE3DTV is more relevant to the artefacts specific to mobile stereo-video compression and transmission. The user experience information gathered offers more benefits to service providers rather than users.

The **iNEM4U** project (Interactive Networked Experiences in Multimedia for You) (iNEM4U Consortium,2009) (Hesselman et al., 2009) is developing an open distributed software framework that allows users and service providers to seamlessly combine interactive multimedia content and services from different types of networks, such as the Internet, in-home, mobile, and IPTV networks, into one shared experience. In iNEM4U, players are allowed to share selected parts of their owner's personal situation including disclosing information on a person's current activity, e.g., whether someone is reading, walking, or sitting down, or information about someone's emotions, e.g., level of anxiety. This ultimately leads to people at remote locations sharing a similar experience. Although the system is an interactive system, its interaction, however, is mainly concerned with how to express the personal situation and thus pays less attention to user's interaction with the content.

The **ROLE** (Response open learning environment) (ROLE Consortium, 2009) project offers adaptivity and personalisation in terms of content and navigation, and an associated learning environment. Users' learning needs are required as the input to the system. Based upon the learning needs the learning environment elements can be combined to generate new components and functionalities. ROLE is a typical interactive expert system. Ambriola and Notkin (1988) argued that the changing the *action paths* based upon users' profiles, i.e. either novice or expert, could help solve a problem more efficiently. One problem with this interactive design is that for some users, a user's learning needs may not be easily expressed via the user input.

LIVE (Janez and Mladen, 2009) produces in real-time a non-linear multi-stream TV broadcast of sports events that adapt to the interests of the viewers. The innovation of this project lies in the feature that TV consumers' feedback is fed into a control room to guide the production process.

For all the foregoing projects, a user centred approach to task interaction is used. User interaction including interactive tasks and user feedback are taken into account in order to establish a personalised experience. Table 2-1 summarises the types of AV-player used in some characteristic surveyed systems.

Table 2-1 User Tasks, Personalisation in A-V Players

A-V Players	**Basic A-V Control**	**Advanced A-V Control**	**Video Content Selection Mode**	**Multi-View Support**	**Passive Adaptive Task**	**Active Adaptive Task**
NoTube	YES	NO	Content	NO	Video Recommendation based upon either friends recommendation or content viewing statistics or user's selections	NO
MYMEDIA	YES	NO	Content	NO	Video Recommendation based upon video content classifications and content viewing statistics	NO
MOBILE3DTV	YES	NO	N/A	NO	NO	NO
iNEM4U	YES	NO	N/A	NO	Notify users with TV channels	NO
ROLE	YES	NO	Content	NO	Offer user required video material via user input query	NO
LIVE	YES	NO	Channel	NO	Centralised live broadcast based upon user's ratings	NO

2.2 User Modelling including Interaction and Task Modelling

One of the core research objectives in this thesis is to model users so that a system can adapt its services to a user's preferences. This section examines existing user models and task models in order to give insights into the use of user modelling within the context of live sports events.

A user or interactive task is user goal oriented e.g. play a video. A task sometimes requires simple interaction, i.e., a single-click or it requires much more complex interaction, i.e., several interactions involving drag-and-drop, text input, etc. In an A-V player, basic controls activate simple user tasks such as play and pause, requiring less interaction effort. More advanced user tasks such as camera switching that were proposed in the My eDirector project initial user survey (My-e-Director 2012, 2008) are crucial in the next generation A-V player but require substantially more user interaction and if user-centred, i.e., personalised, can substantially increase the operational load on users.

An interactive system by its nature involves human and computer interaction. Mainstream definitions of HCI (Hewett, 1992; Dix et al., 1993; Preece, 1995) contend that HCI mainly concerns human factors, computer system usability and explicit interaction between human and computer system. Whereas, implicit HCI (IHCI), a concept first proposed by Schmidt (2000) is formally defined as "*an action, performed by the user that is not primarily aimed to interact with a computerised system but which such a system understands as input*"(Poslad, 2009). Most explicit HCI research work tends to focus on user-centred requirements and study how computer systems perform tasks to meet users' requirements (a mental modal). Implicit HCI, in contrast, is more concerned with systems model users and their actions.

The AI community has different views of how human and computer should interact (Winograd 2006). More recently, the term interactionist AI (IAI) proposed by Leahu et al. (2008) somewhat harmonises both camps. It is defined as: *the concrete, technically feasible AI approaches for supporting real-time, intelligent interaction with a changing environment.* This definition completes the IHCI definition in terms of expected system actions after receiving the user input. That is, when an interactive system is designed with both IHCI and IAI methodologies, the interactive system will be able to utilize AI approaches to respond users intelligently. Nevertheless, there is still a remaining problem that the target system should act intelligently may not be clear. In this thesis, the system will target user preferences when a user is viewing live sports events.

Personalisation involves *tailoring applications and services specifically to an individual's needs, interests, and preferences* (Poslad, 2009). When personalisation involves an interactive system, user interaction needs to adapt more to user preferences. Hence, personalised interactive systems are similar to an adaptive system (Benyon and Murray, 1993) that is intelligent. The 'intelligence' here suggests that the system can reach its adaptation objectives via learning. Hence, a personalised interactive system in this thesis is defined as: *an intelligent system that can alter aspects of their structure, functionality or interface on the basis of a user model generated from implicit and/or explicit user input, in order to accommodate the differing preferences of individuals or groups of users and the changing preferences of users over time.* Within the context of live sports events, a personalised interactive system is able to assist user interaction, to more intelligently tailor multiple video stream content according to users' possibly changing preferences, as live sports events progress.

In a personalisation application, the user preference seen as the knowledge of users, is critical to the system. A user model at a high level represents a system's knowledge of a user. A user model or user profiles can consist of a collection of a user's personal data including identity, nationality, system usage data etc.

As the nature of personalised interactive systems is to allow a system to adapt to users' preferences, this requires the system to differentiate between users in terms of different individual or group information. Two tasks are critical in building such user model. The first one is the definition of the user information, i.e. what kind of information should be maintained in the user model. The second is the application of the user model, i.e. how to use the user information. Existing approaches for user modelling can be grouped into three categories, cognitive based, empirical based and usage based.

2.2.1 Cognitive and Empirical User Models

A cognitive model concerns a user's internal reasoning strategies. Cognitive psychology plays major role in shaping a user model. When using a cognitive model, users are thought as different individuals (Rich, 1983). User's cognitive capabilities and weaknesses must be taken into account (Vander Veer et al., 1985). The key aspect that characterise a cognitive model varies. In Yallow's work (1980), users' spatial and verbal capabilities are investigated. In experiments, users with low and high spatial capabilities received content in either graphical/spatial or verbal format. The results suggest that the immediate retention of material is better in a format in which a subject has high abilities.

Robertson (1985) in his work found there are several cognitive styles that affect a user's interaction. For example, users have differences in the way they distribute and allocate attention resources. Users also differ in the way they pay attention to different aspects of a task. Users also have differences in planning the strategies to complete a complex task. Because these cognitive styles involve high level processing, it is not clear how they can be exploited in an adaptive interactive system. In personalised interactive systems that focus on individual differences alone, they may face computational pressures on systems when dealing with many users, i.e. scalability challenges, e.g. in a recommendation system.

An empirical model mainly concerns the acquisition of knowledge a user possesses via observations from experiments. Using an understanding of a user's knowledge of a task, a system can make decisions on how to assist users. An empirical model is more dynamic than a cognitive model. A user's knowledge of the system can increase when they can more perform a task proficiently whereas a user's cognitive capabilities often evolve more slowly. As a result, it requires a more dynamic maintenance mechanism to keep empirical user information up to date. A further issue with an empirical model is that it is normally difficult to retrieve user's knowledge explicitly from users themselves as users sometimes may not clearly describe their attitude or knowledge about something, i.e. the user knowledge retrieval challenge

2.2.2 User Task Models

The notation of a 'task' has been central to work in design an interactive system since HCI started (Benyon et al., 2005). In a personalised interactive system, task modelling aims to design a model that can assist users to finish a task more effectively. Existing task modelling techniques mainly model the task from three different aspects, namely the logic of task, the cognitive analysis of the task and the structural knowledge related to the task.

Hierarchical task analysis (HTA) (Annett and Duncan, 1967) represents task structure in a hierarchical structure. User tasks in HTA are structured in different plan paths so that tasks and sub tasks can be logically linked. In Stanton's work (2003), HTA is used to look for error situations using the model. Lim and Long (2009) modified the original HTA model, replacing task goals with object names. It is difficult to use HTA to model interactive tasks as it does not explicitly model task interaction. The GOMS model (Card et al., 1983) is often used to describe the user knowledge needed to perform a task in terms of Goals (i.e. what user intends to accomplish), Operators (i.e. actions performed to achieve the goal), Methods (i.e. sequences of operators), and Selection rules (i.e. rules to select certain methods over the others). It assumes the task model developed can subsume all the methods that accomplish the task. The selection rules must be well-learned sequences of sub-goals and operators (John, 2003). In adaptive interactive systems, GOMS is unable to address the errors made by users as it assumes that users are expert users and always know how to, and are able to, do the right thing at the right time. Entity-Relationship Modelling and Information Artefacts (ERMIA) (Green et al., 1996) is a task analysis model that deals with the descriptions of knowledge structures. Rather than modelling tasks as steps, it describes user knowledge via entity relationships including 1: m, 1:1 and m: m. To a certain extent, ERMIA seems to be focused more on designing the user interface rather than on user tasks; it does not model the relationships between tasks.

In practice, a number of approaches have been used to model user tasks. The User-Task Elicitation Tool (U-TEL) (Tam et al., 1998) is able to elicit user task models. It processes the descriptions of tasks through dissecting input textual information and mapping verbs to tasks and nouns to objects of these tasks. U-TEL was initially designed for domain experts without any knowledge of interface modelling and programming, which makes U-TEL useful but too simple to formally describe general tasks, furthermore the elicited tasks are not structured. The Convenient, Rapid, Interactive Tool for Integrating Quick Usability Evaluations – CRITIQUE (Hudson et al., 1999) was initially designed to produce GOMS (Card et al., 1983) models via analysis of the past event logs, It looks at historical use of systems but does not predict future use. The ConcurTaskTrees Environment (CTTE) (Mori and Santoro, 2002) is a task modelling development environment that uses hierarchical structures. The difference between CTTE and HTA is that CTTE defines more types of tasks including abstract, interactive, human, system ones and etc. In addition, CTTE defines a set of task relationships. IdealXML (Montero et al., 2005) is a modelling tool describes user task models and interfaces based upon the user interface extensible Markup Language (UsiXML). UsiXML is XML based language that is able to describe user tasks and interfaces. As both CCTE and IdelXML use CTTE based notations and XML based descriptive language, they also inherit some limitations from both CTTE notations and XML, i.e. CTTE does not allow representing contexts of use (Trevisan et al., 2004) whereas XML using an arbitrary data structure lacks of formal logic expressions. UsiXML is also no more meaningful than a textual description of CTTE diagrams.

For modelling interactive tasks, CTTE is powerful as it provides an explicit notation for task types. However, when modelling personalised interactive tasks, modelling dependency relationships between tasks is not sufficient. CTTE can describe a sequence of tasks as a work-flow. However, it has no notation to represent whether and when a task can be redone or forwarded. In a personalised interactive task model, new notations are required. It should be able to explicitly describe the context in which the system can adapt to user preferences via traceable user actions.

In most cases, task modelling mainly deals with the task design and is used to leverage the system performance of the task. There is no information about what user is imputed in the system through task modelling, such as how user performs the tasks and what are their preferences. In order to enable the system have the knowledge of user, a user model is required which is investigated in the next section.

2.2.3 Comparison of User Models

A cognitive model is not often reported in the field of computer science as it is more viewed as a subject in the field of cognitive science which draws a lot on the computer system among other disciplines. An empirical model is often used in tutoring systems in which users are often classified by the system as either novice or expert. In practice, an empirical model is often constructed based upon usage information. In Manson and Thomas's adaptive interactive system (1984), the number of times users log on to the system and the types of commands used by users are used to assess the user's expertise level of using the system. A user's expertise level can be upgraded when some threshold value has reached. One problem when using this approach is that a user progresses from novice to expert level not in a step wise manner but in a continuous way; some users may not expect adaptive changes of the system even though they have been interpreted by the system as expert (Norcio and Stanley, 1989). A usage based user model is often used with multimedia content delivery systems for interpreting users' preferences for content. In (Zaletelj et al., 2009), user feedback about real time sports video content and user channel switching information are collected in order to compile a user profile about content viewing preferences, to facilitate real time TV production. TV Recommender systems are a typical application of a usage based user model (Middleton et al., 2004). However, a usage based user model may not be well performed when a systems lacks sufficient explicit user input. Importing secondary user information (i.e. data from third party systems) into the system (Bellekens, et al., 2009) may overcome this problem, however such secondary data may not be available and even when it is available, additional issues such as data interoperability and privacy could arise.

A user's usage model contains a user's interaction history. Usage varies across different tasks. Usage information often cannot differentiate between users but this depends upon the task and the number of users. Usage data often requires an additional tool such as a statistical tool to interpret it. In the domain of a multimedia content delivery system, image and video associated variables such as scenes, frames, descriptive keywords and symbols are taken into account to produce user profiles (Manzato et al., 2009). A usage model relies on a user's physical interaction with the system. To an extent, a usage model by its very nature can obtain more up to date user information. However, using the right approach to compile and interpret this information thus becomes a critical task to obtain a reliable user preference in a personalised interactive system, i.e. usage information interpretation challenge.

2.3 Personalised Viewing of Sports Events

The main factors that affect the process of personalisation such as content acquisition, user preference retrieval and type of adaptation method, are surveyed in this section.

2.3.1 Content Acquisition and Production

Extensive research has been devoted to the personalisation of broadcast sports events. This is in part driven by commercial opportunities to potentially raise revenue through offering value-added personalisation services. Much existing work on personalisation tends to be based upon video content acquisition while other work is more concerned with broadcast production. These involve sports video content indexing, retrieval and summarization and personalising viewing. The more structured and meaningfully annotated that content (features) can be produced, the greater the potential for more finely grained personalised interaction with parts of the content rather than just with the content as a whole.

2.3.1.1 Sports Video Content Low-Level Feature Based Personalisation

Studies of low-level and medium-level video content features usually include scene and shot detection, object tracking, content descriptor extraction and matching and semantic document analysis. Shot and scene detection and object tracking can be employed to personalise the viewing of the events in terms of scenes/shots or incidents.

Li et al. (2009) use a bag of visual words model to represent the key frame for each shot and to classify these shots against predefined shot types. In Tan et al.'s work (2000), camera motion is detected and classified in terms of wide-angle and close-up shots in basketball video and to detect events such as fast breaks, full court advances and shots at the basket. A different use of classified shots can be found in Kameda et al.'s work (2004). In their prototype system, a free view system is built to allow user to view a soccer event from arbitrary viewpoints. This is done via the reconstruction of 3D environments via extracted video the textures and player positions from many pre-set cameras. Sports objects such as athletes can also be tracked to enrich a new viewing experience, this can be helpful when there is less camera coverage in some sports events, e.g. in Hallberg et al.'s work (2004) the skiers' positions are tracked and presented in a user end terminal.

To achieve sports video indexing, retrieval and summarization and hence to enable personalisation services, the annotation of the video content is required beforehand. Existing work on sports video annotation can be classified into structure based annotation, incident based annotation and ontology based annotation. Structure based approach annotates the video with respect to video's metadata about frames or shots and sports information such as plays and breaks. In Ekin et al.'s work (2003), the frame and shot type information are used to annotate the video. Xie et al. (2003) segmented the frames by labelling them as plays and breaks in soccer events. Incident based annotation relies on the analysis of audio, visual and textual of a sports video content. Miyauchi et al. (2002) detected the semantic incidents from broadcast sports video by analysing the audio energy and frequency spectrum. Li et al. (2010) audio stream segments are classified to obtain the cheering events which can be further used to detect the location of highlights. In Katsarakis and Pnevmatikakis (2009)'s work, camera motion is detected to identify the sports incidents in a long jump event. Caption texts overlaid on sports video is another resource from which the incidents can be detected. This can be found in Zhang et al.'s work (2002) and Assfalg et al.'s work (2003). In Xu et al.'s work, web-casting is used to intuitively detect events by finding the sentences including the relevant keywords. Ontology based annotation approaches are able to annotate the video content more semantically (Bertini et al., 2007). For example in Strintzis et al.'s work (2004) a Visual Descriptors Ontology and a Multimedia Structure Ontology based on MPEG-7 Visual Descriptors were used to support content annotation.

Despite the fact that the foregoing approaches can identify different aspects of the video content, such as sports incidents, the limitations are also obvious. Firstly, most of this research has focused on pre-recorded videos rather than live feed, for example, the video retrieval system and summarisation systems. Secondly, some of the approaches are only applicable to certain sports events types, which make them less flexible to apply to different sports. Thirdly, this research tends to over emphasize on the low-level features of the video content, e.g. some general video metadata such as 'football'. This may not allow a direct matching between the sports video content and more complex and multi-valued user preferences. One example is that if a user's preference is expressed as 'a team sports in a stadium, this indicates that football is one of the preferred events. However, as the keyword 'football' is not expressed it will be difficult to directly match the video metadata 'football' to this preference.

2.3.1.2 Broadcast Production Process Based Personalisation

Instead of studies that focus only on the video content itself personalisation in the video production process is also being investigated, i.e. to introduce personalisation services to the broadcast production process. As a means to achieve these personalisation services, allowing user to interact with the video content is essential. Wage et al. (2006) in their work argued that, mass media events such as sport events are the perfect 'playground' for the introduction of interactive formats. In the FP6 EU project IST-LIVE (http://www.ist-live.org), human directors are the target users of the personalisation services, in this project multiple video streams are available in the human director's control room and human director is allowed to annotate the content and to filter the live feeds. In contrast to IST-LIVE, it was argued that audience' preferences should be considered in broadcast production (Wages et al., 2006). In coincidence to this, personalisation has been proposed to allow end users to participate in the live show production, e.g. in FP7 EU projects My-e-Director 2012 (http://www.myedirector2012.eu), users are allowed to direct their own live sports events.

Personalisation based upon video content is the most used personalisation approach at present. It relies on low-level or medium-level video features analysis, and hence it by nature is less user-concerned and can be less suitable in live broadcast scenarios. Personalisation based upon broadcast production is becoming popular as it could potentially add more enhancements and interactive options to end users, and is designed to be applied to live sports events broadcast; however the challenge of doing so is to retrieve target users' preferences. In the next section, approaches for user preference elicitation issues are reviewed.

2.3.2 User Model Acquisition: Preferences

User profiling was initially applied to Internet applications (Claypool et al., 2001) to acquire user preferences. The user preference elicitation approaches are mainly determined by the understanding of the preferences in terms of its 'volatility', i.e. dynamic preferences versus static preferences.

In Sung et al.'s work (2002), a user preference is defined as a function on how a user likes a given item based upon the properties of the item. User preferences can be determined by several factors, in Manzato et al.'s work (2009) image and video associated variables such as scenes, frames, descriptive keywords and symbols are taken into account to produce user profile. Despite the fact that user preferences can be multi-faceted, they are more often assumed to be static across different context descriptions (Jembere, Adigun, and Xulu 2007). In literature such as (Buchinger et al., 2009), (Eronen, 2001), (Jumisko-Pyykko et al., 2008), (Shelley et al., 2009), (Menon et al., 2005), (Rice and Alm, 2008), (Svoen, 2007), etc., user's attitude towards new technologies and services are used to imply user's preferences as it is believed user attitude is relatively stable over time. In stark contrast, it is also believed that user preferences are more dynamic in terms of time among other factors. For example, in Panayiotou et al.' work (2005) and Jung et al.'s work (2002, 2005), user preferences are envisioned to be updated in terms of time. It can also be described temporally such as long-term versus short-term (Sugiyama et al., 2004).

In general, there are two types of user preferences elicitation approaches: an explicit approach and an implicit approach. Explicit approaches normally ask the users for their preferences via a graphical user interface such as a form or an optional menu. For implicit approaches, user preferences can be retrieved from user interactions with the system (Middleton et al., 2004). Implicit approaches can often be found in existing TV programme recommendation systems, in which usage logs are used. In Zaletelj et al.'s work (2009), user feedback of the real time sports video content and user channel switching information are collected to model the viewing preferences in order to facilitate the real time TV production. The advantage of using implicit model is that it is unobtrusive to user's goal in using the system (Martinez et al., 2009) (Sarah et al., 2008) (Youngblood et al., 2005), the main drawback of this approach is that lack of user input may bring the user to confusing or frustrating situations (Sarah et al., 2008). For the explicit model, it is argued that it provides more accurate data describing user preference in Web-based information retrieval applications such as in Martinez et al.'s work (2009) However, system relying on user input of preference information may result in minimal sets of user preferences retrieval (Yoshihama et al., 2003) (Sousa et al., 2006). Studies in (Faltings et al., 2004) found people are often unable to state their preferences beforehand. The explicit model in some video rating systems may also make preferences less explicit. For example, in systems such as You Tube, it is difficult to tell what user preferences really are in terms of scalable star ratings, this is because there is neither an objective rating standard from the user side nor an explicit preference definition from the system side. In the problem domain of live sport events broadcast, user preferences can be rather dynamic as it can be affected by many additional factors such as available sports on air, available cameras and incidents and etc. Due to the dynamic nature of live broadcast, the system may need to understand the user preferences as the live show progresses. An explicit approach is not impossible; however, the cost can be high as user may need to input the preferences from time to time. Using implicit approaches can completely avoid the shortcoming of explicit approaches as the system can obtain user preferences via monitoring the interactions. Nevertheless, problems could still occur. For example, one problem is the cold-start problem (Maltz and Ehrlich, 1995) which occurs when a system lacks initial explicit user input. Bellekens, et al. (2009) suggested a solution by importing secondary data of user information to the system, but this only works when secondary data are available.

Live sports events scenarios such as Olympic Games are different from most video-on-demand scenarios because user preferences may often change as events progress. In other words, the user preferences can be highly dynamic. A personalised interaction system in these scenarios is required to understand user's preference in timely order to tailor services that match the dynamic user preferences. User preferences can be either implicitly or explicitly retrieved, an implicit approach can be preferable when user preferences updates more frequently whereas an explicit approach can be helpful in dealing with user preferences that are relatively static across use sessions.

2.3.3 User Acquisition Model, Interactions and Tasks

In an interactive system, a major means to acquire user's preferences is via acquisition and analysis of user interaction. Hence the questions of: 'which approaches (i.e. implicit vs. explicit) to use and what kind of user information of user preference should be collected' are of importance. In the scenario of viewing the live sports events, there are three kinds of interaction. These interactions include the interaction with (sports) events selection, interaction during viewing events and the interaction to replay the scenes of an event.

2.3.3.1 Interaction with Sports Events Selection

This interaction requires the user to express their preferences beforehand. The system needs to understand the individual user's preferences of sports events with respect to multi-faceted properties such as event type, event venue and athlete's performance. At present, explicit approaches are often used. In Schalleck et al. (2004) and Van Beusekom et al.'s work (2004), a score voting system is used to allow users to express their own judgements concerning athletes' performance. Zhang et al. (2007) explicitly asked users to express preferences for pre-prepared video segments. Ren and Jose (2006) used attention analysis to model users including directors, the audience and commentators, and to detect the event highlights. One typical application of the retrieved user preferences is the recommendation system which offers tailored recommendation to the user by analysing the user's preferences.

Although different techniques can be used to retrieve user's preferences, existing approaches are applied mainly in an explicit manner. This approach may work properly in video on demand system in which videos are well annotated. In the live sports events, user may switch between events more frequently because of new upcoming events and many concurrently occurring events. A system featured with recommendation may need to recommend users the events frequently. With explicit approach this means users will be asked before each recommendation. Hence the challenge here will turn out to be how to balance the trade-off between obtaining accurate user preferences as events progress and reducing user-to-system interaction complexity.

2.3.3.2 Interaction during Viewing Events

Interaction during viewing the events enables user to interact with the video content and perhaps direct the broadcast of the live sports events. Intuitively, the interaction a user could do during the viewing can be either to interact with the objects within the videos or interact with the viewing angles directly. As of today, there is little published about how to retrieve user preference from these interactions, much of the work tend to focus on interactive technique studies.

2.3.3.2.1 Interacting with Objects within Videos

With regard to the objects of interest, the interaction will allow users to locate the objects of interest so that they can be tracked by users. In general, there are two approaches to achieve this. One is to label the objects and the other is to draw user's attention to the objects via zoomable user interface.

Most labelling approaches rely on the video analysis such as background colour extraction (Cheung et al., 2000) multi-view stereo set (Seitz et al., 2006) etc. In Li et al.'s work (2008), the multi-object silhouette cues are used to represent the 3D shapes of the objects. House et al. (1998) proposed a framework to assist users to label objects based upon the analysis of a sequence of extracted images. The well-known problems of using the first approach somehow make it difficult to extend this from lab trials to real scenarios. For example, the colour extraction may be inaccurate when in an outdoor environment and when environmental brightness changes. Multi-view stereo set may also be less effective when there are some regions not covered by the pre-set cameras. In addition, in order to identify the labelled objects, the video stream metadata are required to be rich enough to describe each visible objects for each frame, this is especially challenging in live broadcast scenarios. Instead of using video analysis techniques, put a sensor on athletes' body can also be an option. For example, Foina et al.'s (2010) used RFID technique to track the athletes to assist coaches to analyse the athletes. Although it appears to be promising, it is still questionable whether it is practical in real scenario, such as it is not clear whether the athletes are allowed or willing to wear such device during competition.

The zoomable user interface (ZUI) approach can be found in most multi-media applications. Zooming on live video content is a function can be found in some commercial (e.g. Apple QuickTime) and open source (e.g. VLC player) desktop video players. For these desktop video applications, the zooming function is triggered by a video player window size change. For small screen device, zooming function is also been studied by using the temporal separated zooming technique as in (Knoche et al., 2007). Much of research attention has been devoted to the development of ZUI theories and interaction techniques. Cockburn et al. (2008) summarized the ZUI techniques into four types, namely overview + detail (e.g. Google map street view), temporal separated zooming (e.g. Microsoft 3D Virtual Earth), focus + context (e.g. Apple Mac OS X Dock) and cue-based interfaces (e.g. semantic zooming such as windows mobile calendar

(Kosara et al., 2001). According to existing literature, the pros and cons of each type is presented in Table 2-2.

Table 2-2 Pros and cons of ZUI techniques

ZUI Technique	Pros	Cons
Overview + Detail	Suitable for document comprehension (Beard and Walker, 1990) (North and Shneiderman, 2000)	Additional use of screen real estate and suboptimal for dynamic Activities (Baudisch et al., 2002)
Temporal separated zooming	Apply to different application domains (Bederson et al., 1996) (Druin et al., 1997) (Furnas et al., 1998)	Additional operational load for understanding relationship between pre- and post-zoom states (Cockburn et al., 2008) (Plumlee and Ware., 2006)
Focus + Context	Suitable for tasks that involve gaining a rapid overview of the data space (Gutwin, 2002) (Zhai et al., 2003) (Hornbaek et al., 2007)	Visualizations distortion (Nerkrasovski et al., 2006) (Hornbaek et al., 2002)
Cue-Based	Allows semantic depth-of-field technique (Renaud and Eric, 2007) (Kosara et al., 2001)	Modified objects rendering form can introduce proxies for objects not expected (Cockburn et al., 2008)

While the existing ZUI techniques tend to be applied on the static zoomable objects, in the context of live video content, some of them may not be applicable at all and for some, additional disadvantages can be revealed. The focus + context technique (e.g. fisheye) perhaps is not a suitable solution for live video content because a rapid zoom-in or zoom-out could get viewers dizzy besides, visualization distortion (e.g. distorted athlete image could make the athlete unidentifiable) can be a critical threat to the viewing experience. For overview + detail techniques, an additional overview of the live video stream could reduce the available bandwidth which in turn could reduce both streams' visual quality. For temporal separated zooming, dynamic live video content could further increase the cognitive load as it is found only one graphical object is held in visual working memory (Plumlee et al., 2006). The cue-based technique could make the live video content more meaningful, e.g. a zoom-in on a moving athlete could trigger an additional GUI showing the detailed information of that athlete. However, live video content metadata must be rich enough so that the system can parse the content in a frame by frame manner and this can be an extra challenge for live video content processing. Among three available ZUI techniques for the live video content, except for the temporal separated zooming technique, all other techniques require additional resources in order to overcome the associated disadvantages. For overview + detail ZUI, high bandwidth is required. For cue-based ZUI, additional metadata associated with video stream are needed. Therefore, given there is no extra resource available, temporal separated zooming technique becomes a more preferred ZUI solution for video content.

Although the need to use a ZUI is widely accepted, a limited effort has been committed to address additional challenges such as the target location shift problem that occurs in a sequence of transient scenes in which the zooming targets no longer at the pre-zooming position after zoom in, the reduced zoomed visual quality problem which when there is an over magnification of video pixel size. A third challenge relates to the zooming visual presentation is the user cognitive load problem (Plumlee et al., 2006), which occurs in a rapid zooming process (e.g. a sudden zoom in) so that users lose the track of pre and post zooming state of the content.

Interacting with objects within videos relies somewhat on the interactive interface that allows user to track the objects of interest. Existing techniques do not yet seem to address the issues regarding the interactive interface for live sports events. Labelling objects within the video is especially challenging in live broadcasts, while the existing ZUI techniques need to be modified and improved to support such use scenarios. Although little literature has discussed the retrieval of user preferences, it is envisioned that an implicit approach would be more appropriate here as users may frequently interact with the system during the viewing. When a certain amount of zooming interactions are observed by the system, a future zooming region perhaps can be predicted by the system which may in turn to enable an automated zoom in.

2.3.3.2.2 Interacting with Viewing Angles

A live sports viewing system with multi-views potentially offers viewers a richer viewing experience than using a mono-view system, especially if viewers have some control of the camera switching themselves. In the user requirements survey (My-e-Director 2012, 2008), 62% of participants thought the manual camera switching is highly desirable. In principal, user-end camera switching can be characterised in terms of three aspects: camera type (virtual vs. real or physical), visual content processing (required vs. ignorable) and camera switching operator (system handled vs. user handled)

Firstly, real cameras require a higher bandwidth cost especially when more than one camera is streamed concurrently, e.g. a picture in picture view of same sports event. Previous works such as those done by Muntean et al. (2004) and Leu et al. (2009) have contributed to the adaptive multimedia streaming, aiming to maintain a smoother visual quality. Some commercial solutions[1] are also emerging such as IIS Smooth Streaming, Flash Dynamic Streaming and Apple HTTP Adaptive Bitrate Streaming. One common character of existing adaptive streaming technologies is that they are mainly concerned with single stream transmission and have paid less attention to multiple stream adaptation. Rather than using multiple physical cameras to create multi-view system, virtual cameras can also be used in visual content authoring systems to create multiple virtual views. Techniques including panoramic views and object with multiple viewpoints are often used (Injae et al., 2006). There are two main advantages when using virtual camera views. First, visual content from a mono-view can be converted to multi-view visual content. It therefore only requires the bandwidth for streaming one AV stream over a content distribution network (CDN). The other advantage is that it allows users to view the content from different virtual viewpoints. However, due to the nature of visual content authoring system, i.e. mono-view, the original visual content must be visually distorted to achieve a virtual multi-view effect and such distortion may undermine the visual quality and create target identification problems. Further than that, the technique is not quite suitable for live visual content processing as of today.

Secondly, virtual camera views can be synthesised after transforming and combining the images from several fixed cameras (Inamoto and Saito, 2007) (Morhee et al., 2006) (Guan, 2009). In the context of Web TV broadcast, such virtual view synthesis is affected by processing complexity, camera settings flexibility and video streaming bandwidth. Inamoto and Saito (2007) found the processing time increased along with the increasing number of dynamic objects within visual content Specific camera settings are also required in these approaches, e.g. cameras may need to be set up only along one side with view overlaps between them. Multiple camera feeds are required to generate the multi-view effect - this demands a higher bandwidth for the CDN.

Thirdly, camera switching is in practice task-driven. For example, multi-camera views can be mainly system driven within which the system primary evaluates the posture evolution of football players (Zhu et al., 2007; Leo et al., 2009). The advantage of automating camera switching is that no manual workload is needed to switch cameras.

[1] http://learn.iis.net/page.aspx/792/adaptive-streaming-comparison

The downside of this approach is that users have less control of the camera switching. For non-expert system problem domains such as the sports video play system (Inamoto and Saito, 2007) camera switching relies heavily on the user control and this is also true for Web TV broadcast systems such as NBC's Sunday football night extra. In these applications, camera-view options are listed in a simple text menu from which users can choose a preferred view and the chosen view will be played in main screen afterwards. Although user handled camera switching allows the user have more controls of the system, it may also increase user's operational workload on camera switching task, i.e. user may repetitively do the switching in order to timely view every single corner covered by cameras.

Table 2-3 summarizes the six types of user-end camera switching approaches in terms of three fundamental aspects with associated binary options.

Table 2-3 Pros and cons of camera switching approaches

Approach	**Pros**	**Cons**
Virtual Camera	Less bandwidth cost Free viewpoints	Visual content distortion problems
Real Camera	True multiple views	High bandwidth demanding when all feeds are required Viewpoints depending on camera positions
Visual Content Processing	Create processed rich visual content	High processing complexity High bandwidth cost
Non Visual Content Processing	No processing complexity Not all feeds are concurrently required	Multi-view depends on number of real camera views
System Handled	Less user operational workload on camera switching	User has less control of camera viewpoints Suitable for expert users
User Handled	User directly controls camera viewpoints More target users	User operational workload may increase linearly while increasing the number of switches

Camera switching in user terminals is becoming popular on Web platforms due to the availability of rich Internet applications. Existing broadcast applications, to an extent, have been able to offer users options to switch to alternate cameras online in live mode. Sometimes multiple views are viewed concurrently, e.g., Picture-in-Picture, in order to enable users to conveniently click and switch between a main view and multiple side-views, possibly using high definition video content leading to have a higher bandwidth CDN provision cost. E.g. Lou and et al.'s (2005) lab system has to encode the multiple camera streams in a small frame size (i.e. low definition images) in order to stream them smoothly over Internet. Whereas the commercial applications such as the NBC's Sunday football night extra allows users to view the content in high definition, however only one camera stream can be viewed at a time. As for the interaction involved in this task, switching multiple viewing angles is a type of interaction can be performed frequently. This is because user may prefer to have a timely viewing of all available angles or unanticipated incidents as a live event progresses. Hence, the personalisation of such interaction will require the system to timely acquire the user preference and thus the implicit approach can be more favourable than explicit approach.

2.3.3.3 Interacting with Replay

Interaction to replay the scenes of an event is different from foregoing interaction types. This is because this interaction by nature is designed for experts who are supposed to use this interaction to edit the videos and serve others. If this interaction can be performed by individual end users then they will have more chances to review the exciting moments or do other editing work on scenes. In effect, allowing end users to do the editing on the time-shift viewing of the live sports events is essential and this should be one of the important characteristics of the next generation A-V player. Existing approaches of editing video content can be roughly classified as either real time editing or post editing. The former can is often used in live broadcast while the latter is more used for multimedia content analysis.

Post editing is a field that attracts much of the multimedia research efforts. One common focus of most of the research is on the extraction of highlights based upon various broadcast events analysis. As in Lee et al.'s work (2007), the commercial event can be detected. Likewise, in Huang et al.'s work (2007) and Han et al.'s work (2009), the replay event becomes focus. In terms of the problem domain, some of the research also studied on particular sports events, in Singh et al.'s work (2006) cricket game is targeted and in Wang et al.'s work (2004) the soccer game is focused. As the post editing is mostly done via the analysis of mixed video content, the feasibility of existing post editing approaches can be restricted. For example, if the replay and live transition pattern (e.g. some transition effects in live sports broadcast) changes or sports event changes, the performance of most replay detection algorithms will be undermined.

Sports video content processing differs from other video content processing partly due to its massive objects content nature. Among existing image processing techniques, motion based approaches have been often investigated. In Ma and Zhang's work (2003), motion texture is used to detect motion pattern. In Xiong et al.'s work (2003) MPEG-7 motion activity descriptor intensity has been used to generate sports game highlights. In Wang et al.'s work (2004), motion vector field information is used to detect sports game camera motion. Although these approaches vary in terms of perspectives of motion factor analysis, the general structure is a general one. That is: a) the low level video motion feature is normally extracted and b)additional processing layers are often introduced to either semantically interpret the low level video features or do some semantic classification of the low level motion data, c) the result eventually is used for a particular problem domain. One advantage of using this structure is that the application can be tailored to the problem domain. However such advantage sometimes can be a downside especially when the middle layer is over 'domain focused'. For instance, in Wang et al.'s work (2004), the system will not be able to deal with other sports events as the motion is semantically defined in terms of the football game athletic actions.

Among real-time approaches, human takes much of the workload. For example, in a typical directing scenario, the director needs to view different types of images from different camera sources, he also needs to issue orders to the vision mixer and determine which replay goes on air. Some editing tools such as RCE[2] allows the director to edit the live sports events without vision mixer it still requires director's expertise to determine the highlights. Devices such as a DVR (digital video recorder) allow terminal users to manipulate the replays, however, the users' experience can be frustrating (Darnell, 2007) due to lack of control and replay precision. User preferences within this type of interaction are different from other types because they are directly impacted by the scenes and it would also be difficult to ask users to express their preferences of scenes in advance. Hence, rather than focusing on the unpredictable user preference, the personalisation here should be providing users with content that may interest users according to the content itself.

2.3.4 Adapting User Content to User Models

User preferences need to be understood by the system in order for the system to actively adapt to them. However, a mere understanding is not enough to allow the system to adapt to these preferences. A personalised interactive system requires additional mechanisms to adapt personalised services for the user's preferences. In this thesis, two types of adaptation are concerned, they are: passive adaptation and active adaptation.

---

[2] http://code.msdn.microsoft.com/RCE

2.3.4.1 Passive Adaptation (Recommender Systems)

In passive adaptation systems, explicit human user input is needed to decide to do the adaptation. In an active adaptation system, the system does the adaptation using previously acquired user input such as user preference information or using implicit user input (see below). One of the most common types of passive adaptation is a recommender system.

A recommender system is able to generate meaningful recommendations to a collection of users (Prem et al., 2010). Recommender systems can be classified as content based and collaborative based. Content based systems model the link between service content and a person's preferences whereas collaborative systems model the link between a person's preferences and other persons' preferences for the given service content (De Vel et al., 1998). Content based systems could more accurately tailor the service to users, providing them with an appropriate amount of user preference data. Collaborative systems can be useful when there are not sufficient data to describe users' preferences such as in cold-start problems (Maltz and Ehrlich, 1995). Approaches to recommender systems can also be categorised as memory based and model based. The former operates over the entire data to make predictions and latter uses the historical data to build models which will then be used for predictions (Tong et al., 2002). It is reasonable to argue that memory approach can be more efficient when used in content based systems whereas the model approach is more powerful in collaborative systems. It might be also true that the amount and accuracy of the user data determine the efficiency of the recommendation system given the appropriate prediction algorithms provided.

In the video relevant problem domains, content recommendation systems are in mainstream use enabling consumers to filter video content to match personal preferences. In practice, in order to scale up matching content to many individual users, content is matched to group user preferences that individuals belong to. An aggregation function is often used to generate recommendations to a group of users in terms of Jameson and Smyth's work (2007) and Yu et al.'s work (2006). In their works, the preference aggregation approaches can be specifically summarized as: 1) merging of sets of recommendations; 2) aggregation of individuals' ratings for particular items; and 3) construction of group preference models. The first approach could reduce the incentive for group recommendation as it needs to obtain individuals' recommendations beforehand. The second approach could easily generate partial recommendations due to a user's subjective rating standard. The last approach, as in Boratto et al.'s work (2009) and Jameson and Smyth's work (2007), aggregates the preferences of individual group members to form a model of the preferences of the group as a whole. This approach can effectively avoid repetitive computations for users for the same recommended items. However, the unreliable user preferences problem still exists when user preferences are ill-defined and the constrained group classification problem could also occur when both individual user preferences and recommended items definition are derived from users' feedback.

In a live sports events broadcast scenario, recommendations should timely reflect the user's preferences which can be complex and updated as live events progress. Content based recommendation can be more accurately reflect the user's preference but the cost will be increasing computations on user front. The collaborative approach is able to overcome this problem by move the computation from front end to back end. However, the challenges still exist such as: how to timely maintain and update user groups who share same preferences within an event viewing session; how to ensure the recommendation is scalable when user number and events numbers increase.

Existing user grouping approaches largely rely on individual preferences which are usually collected via either an explicit or an implicit approach or both. Explicit preferences can be retrieved via explicit user interface as user preference input (Jameson et al., 2004). Implicit preferences are often obtained via monitoring user interaction with the system (De Ávila and Zorzo, 2009). As these approaches link user preferences directly to the user group definition, user group re-clustering overhead cannot easily be reduced during live sports events scenarios.

Rating systems are often used to present users' preferences about sports events (Boratto et al., 2009; Masthoff and Gatt, 2006; Xu et al., 2002) However, it is normally difficult to calibrate these subjective ratings among a group of users. Consequently, preferences among users will be inconsistent and incomparable. User personality (i.e. cognitive model) can be used as the determinant of the preferences (Recio-Garcia et al. 2009). The weakness of this approach is that it may require the system to study each individual first; the system is less scalable as the number of users increase.

Recommended item tags and metadata are used for recommendation generation (Hölbling et al., 2010). One constraint of this approach is that it requires a large amount of user input tags that can be associated with recommended items. Such user preference modelling approaches are very much user feedback centred and rarely support more finely-grained multi-dimensional characterisation of domain objects, hence recommendations tend be very coarse grained, e.g., athletics or even running rather than the 100m.

2.3.4.2 Active Adaptation

Active adaptation is where the system adapts a service, e.g., content delivery, without explicit input at the time the adaptation occurs - the adaptation is automated. This active adaptation to a user needs to take into a model of the user or user profile. Research concerns the detection of user actions associated with human computer interaction. Fischer and Nurnberger (2008) created adaptive system architecture for the use in vehicles via multimodal interface using visual information and speech as output channels, and manual input and speech as input channels. Rodriguez et al. (2008) allowed the body gesture to control the interaction with an intelligent cash machine. Granic and Glavinic (2005) used a set of user behaviour rules to infer the system adaptive behaviours in a computer-based education system. Nakajima and Satoh (2005) used a spontaneous interaction approach in which the devices in a home environment would automatically detect the users and maintained the user's information for personalisation services. Francois et al. (2009) detected the interaction styles and patterns to allow an adaptive robot to play with children with autism. Despite the fact that all these approaches did allow a system to recognise the pre-defined user actions, they seem to be less concerned about the interaction complexity issue and more concerned about the adaptive system response, in other words, most of them do not use the recognised user actions for further reducing the user interaction overload. In addition, some critical issues are not addressed in these works such as when such automation should be performed and how confident the system is when performing the adaptive actions.

It is still somewhat unclear how to use system automation to reduce the user interaction workload. User action detection is envisioned to be necessary to achieve this though most of the work tends to study the user actions only for the purpose of producing some adaptive system actions. To support such system automation, system actions quality should also be ensured so that the system's takeover-actions better matches the user preferences.

2.3.4.3 Adaptation Methods

User model adaptation normally can be achieved by an association function. Such association functions can vary in terms of the problem domains and proposed problem solving models. But in general, data mining and machine learning techniques are often used. Data mining is the extraction of sequential patterns (Srikant and Agrawal, 1995; Mannila et al., 1997) which analyses temporally ordered data in order to model repetitive behaviour. Machine learning is about finding and describing structural patterns in data.

Data mining usually involves a practical implementation of machine learning techniques, finding and describing structural patterns in data as a tool for helping to explain that data and making predictions from it. Kim et al. (2004) combined statistics and user actions to create their predicted user preferences. A similar approach was used by TiVo system (Ali and Stam 2004). In addition to be used in the association function, machine learning techniques can also be used to validate such association. Ardissono et al. (2004) used two different methods, Bayesian statistics and Decision Tree to evaluate an implicit recommender system. The results showed that fusing the results of the two methods improved the robustness of the system. Other techniques such as SVMs (Support vector machines) (Joachims, 2002) and neural networks (Nichols, 1997) have also been used to explore the implicit feedback. Probabilistic model is also often used as it can be used as a predictive model. For example, Bayesian networks, which use graphical models to specify the conditional probabilistic dependencies between different variables, are increasingly studied for the purpose of discovering some hidden user factors from a sequence of resulting data such as in (Patterson et al., 2003; Patterson et al., 2004). Markov models and Bayesian networks are investigated in several applications, e.g. (Burghardt and Kirste, 2007) used the Markov model to infer user intentions from sensor data. (Aggarwal et al., 2007) proposed a model to identify fragments of the strings based upon hidden Markov models. Although these models achieved their design objectives, they rely on the topology of the symbols in a Markov model, and the design of the topology is often a matter of skill and experience (Aggarwal et al., 2007). Despite the fact that data mining and machine learning techniques are often used as predictive models which are critical in personalisation service, they must be fine-tuned to match the different personalisation models.

2.4 Personalisation Evaluation

One of the main objectives in this thesis is to propose an evaluation framework for personalised interaction. This section examines the existing evaluation approaches.

One fundamental goal of the use of personalised interaction system is to make the complex interaction easier. However, it could mislead users' understanding of the system when the personalisation does not perform as well as it was intended, e.g. recommend an item that the user does not like to view. In order to address this issue, an effective personalisation evaluation method is required. Such evaluation is envisioned to serve different purposes including service quality verification, problem detection and decision making support (Jong et al., 1997) (Van Velsen et al., 2008). Unfortunately, the evaluation of personalisation is neglected in most existing cases.

Studies in (Van Velsen et al., 2008) (Soui et al., 2008) (Spiliopoulou, 2000) and (Lawrence, 2001) revealed the difficulties of conducting an evaluation on personalisation system and found there was no existing standard evaluation method adequate for personalisation. In (Akoumianakis et al., 2001) (Van Velsen et al., 2008), it is argued that the criterion for assessing technology appropriateness is whether it can adapt to the user or a group of users and their context. The argument epitomizes the challenges associated with personalisation system evaluation – i.e. what does adaptation really means for a personalisation system? Their view is that three major factors constitute these challenges. First, personalisation is generally user centred and this requires the system to be adaptive to different users input. Due to the fact that users vary in terms of various complex factors such as culture, habits, these factors cannot easily be captured and are not captured by a system. Second, a user model is normally used to underpin personalisation. However, such user models differ with respect to different system problem domains, i.e. each user model only adapts to targeted contexts, which means the prospected evaluation method is domain bounded. Third, service output adapts to different user preferences, which makes evaluation challenging because of lack of service quality standard.

This section gives a review on the state of the art of feasible evaluation approaches, they are: comparison based approach, knowledge based approach and Hypothesis Testing based approach.

2.4.1 Comparison based Approach

In (Chin, 2001), the empirical evaluation is defined as an appraisal of a theory by observation in experiments. The key processes in an evaluation can be summarized as comparisons and observations.

Table 2-4 Comparison based evaluation designs

Comparisons Design	Expected Result Statement
Personalisation Vs. Non-Personalisation	System with personalisation is better than with system with non-personalisation
Personalisation Vs. Personalisation	System with personalisation technique A is better than system with personalisation technique B

Evaluation using comparisons based upon the classic experiment design theory such as in (Campbell et al., 1966). Such an approach is routinely used in a wide range of scientific fields and most traditional computer systems. In (Van Velsen et al., 2008), it is found that 22% of the personalisation system studies use comparisons as the evaluation approach. Table 2-4 shows the comparison designs and expected meaningful statements of the evaluation.

For empirical evaluations with comparisons, a control group and a treatment group of randomly picked users are usually involved. Both groups will be measured in terms of a desired metric. This approach although is formal and widely accepted, its requirements of comparison setting could introduce several challenges to evaluate a personalisation system.

Firstly, such a comparison experiment is ideally suitable in natural settings. As a result, it will be difficult for business entities and research institutes to conduct such experiment. For the former, it can be a commercial loss to organize a control group with no personalisation for a long time. For the latter, it can be expensive to conduct an experiment in natural settings.

Secondly, identifying what to be compared can be a problem. If the comparison is between a system with personalisation and one where personalisation has been removed, it could possibly lead to the false comparison as the system without personalisation is no longer a worthy opponent (Höök, 1997).

Thirdly, conclusions concerning personalisation quality can be an issue. It is vague to just say one system with personalisation is better than another with or without personalisation. In (Alpert et al., 2007) (Van Velsen et al., 2008), it is argued that most existing comparisons say little about aspects such as perceived output.

There are no specific rules concerning the design metrics used in comparison experiments. For example, an expected evaluation result of "improved user experience" may eventually relate to a number of interwoven user aspects. Though not conclusive, the existing literature shows that user performance is not recommended to evaluate the quality of the personalisation. It is argued some users may have better experience with traditional system than personalisation system (Van Velsen et al., 2008), moreover, the results of user performance can be affected by the other system aspects such as UI layout and designed user tasks.

2.4.2 Knowledge based Approach

In general, knowledge based approaches rely on prior knowledge on what is expected to be the outcome of the personalisation or the purpose of personalisation. If a commercial Web site such as Amazon.com uses a particular personalisation system, then to increase revenue will be the expected outcome of the personalisation. Therefore, the expected evaluation results based upon this prior knowledge will be: 'if the personalisation is good then the revenue increases'.

Knowledge of personalisation results can be implied from the motivation of personalisation. However, such knowledge should be specific enough to distinguish whether the parameters that are expected to be influenced by the personalisation or not.

The evaluation of personalisation thus boils down to assessing the parameters that influence it and controlling the irrelevant parameters in evaluation process. E.g. evaluate the total increase of the income of website for the last three months. Here 'increase of income' is the influenced parameter whereas the 'last three months' is the controlled temporal parameter, which means months before that are not examined. Table 2-5 summarizes the general design of knowledge based evaluation in terms of knowledge design and use of knowledge.

Table 2-5 Knowledge based evaluation designs

Knowledge Design	Use of Knowledge
The parameters that are envisioned to be influenced by the personalisation	The parameters are evaluated and examined against the expected results during the evaluation process.
The parameters are not concerned with or, are irrelevant or the effect of personalisation is uncertain	Parameters are intentionally minimized or controlled during the evaluation process.

Knowledge based evaluation approach gives examiners less constraints and clear evaluation metrics in comparison to the comparison based approach. However, some issues regarding to the empirical experiment designs could still pose challenges to evaluate a personalisation system.

First, the evaluation can be inaccurate due to lack of comprehensive understanding of knowledge. In order to get a comprehensive prior knowledge, some hidden relevant parameters, which may affect the recognized relevant parameters, also need to be considered. For example, change of the users' disposable income (hidden parameter) could directly impact the website income (relevant parameter). User centred data collection methods can be a tool to discover the hidden knowledge. Existing most used research methods include qualitative questionnaire, interviews, focus group, think-aloud protocol and expert reviews (Van Velsen et al., 2008). Although they are useful, the drawbacks of conducting such research are also obvious. For one thing, it can be a resource consuming process, for another it may require an iterative process (Vuolle et al., 2008) in order to make the collected knowledge effective.

Second, the irrelevant parameter may need to be properly controlled. As a result, the laboratory based setting is preferred than real setting, which, to an extent, contradicts to the application of personalisation in the pervasive computing field. Also in (Akoumianakis et al., 2001), it is argued that laboratory observations are not ideal when the social context is taken into account in a personalisation system.

The nature of knowledge based approach requires the evaluation setting to be controlled which undermines the application of this type of evaluation in a real setting. While this approach can be feasible for a pure lab based testing, but the obtained results can be hardly imply the system's performance in real settings.

2.4.3 Hypothesis Testing Based Approach

While comparison based and knowledge based evaluation approaches are not appropriate for the assessing personalisation systems, an alternative approach is required to address the challenges. Table 2-6 shows a summary of how challenges are addressed by existing approaches and should be addressed in any prospective alternative approach. The last column of the table shows the expected prospective alterative with its solutions to the evaluation challenges. In general, the prospective alternative evaluation approach should be able to provide a baseline for different personalisation systems evaluation and be open enough to accept multiple solutions to the challenges.

Table 2-6 Evaluation approach and solutions to challenges

Evaluation Challenges	**Comparison based approach**	**Knowledge based approach**	**Prospective Alternative**
User testing population	Must use 2 or more groups	Group number is not strictly required	Group number is not strictly required
Evaluation Setting	Real setting	Laboratory setting	Accept both settings
Metrics	Appropriateness not guaranteed	Appropriateness can be partially verified via prior knowledge	Appropriateness can always be verified

One candidate alternative can be the Hypothesis Testing based approach. A statistical hypothesis is a statement about the parameters of one or more populations (Montgomery and Runger, 2010). The objective of the testing is to produce a statistically significant decision on whether the hypothesis is merely correct by chance or not.

The core technique of Hypothesis Testing is based upon the use of statistics to determine the probability that a given null hypothesis is true. This effectively gets round the limitations imposed by comparison and knowledge based approaches. For user testing population, control group will not be a mandatory requirement as a single group of users will be enough to prove the personalisation system has affect the users given certain hypothesis is true. For evaluation settings, both real and laboratory settings can be supported by Hypothesis Testing. This is due to the fact that the testing is parameter statistics specific by default so that no extra control is mandatorily required. For metrics appropriateness, the Hypothesis Testing assesses whether the metrics are statistically correct or not, therefore will always verify the appropriateness of the metrics. The basic steps to conduct a Hypothesis Testing can be summarized as shown in Table 2-7:

Table 2-7 Hypothesis testing steps

1. Identify the parameter of interest in a particular problem domain.
2. Propose the hypothesis H0
3. Choose significance level α
4. Determine test statistic
5. Determine the rejection region for the statistic
6. Decide whether or not H0 should be rejected and report that in the problem context

A Hypothesis Testing based approach can be used either as a support to knowledge based approach given that the drawbacks of the knowledge based approach do not concern the personalisation examiner or as a standalone testing method. With a supportive role, Hypothesis Testing can be used to mitigate the interference produced by real setting parameters that can be controlled by laboratory settings. With an independent role, Hypothesis Testing can be used to evaluate the quality of personalisation system in terms of expected hypothesis on testing user populations.

2.5 Summary

Existing research projects tend to investigate the video retrieval system which allows users to pick up or share preferred content. User preferences are mainly determined by content catalogue description and user text input, which can be less reliable for live video content as they change as events progress. Moreover, the interaction in much of the work is mainly used as a content retrieving tool rather than allowing users to interact with the content. As a result, a more advanced interactive video terminal system that is able to capture reliable user preferences that change in terms of live video content; match the content to user's preferences, allow users to interact with the video content; and assist users with interactions is required. In this wok, such system is termed as a personalised interactive system.

Different aspects of a personalised interactive system are reviewed. Issues are identified which indicate that the existing approaches seem to be unable to support a personalised interactive system for live sports event viewing. These known issues somehow give a guideline of the design of a personalised interactive system for live sports events viewing. The found issues of existing work can be summarised as:

1. Existing A-V players offer limited advanced interactions in the context of live sports events broadcasts and lack of personalisation features that allow a player to actively adapt to users preferences.
2. Existing task models are not able to describe interactive tasks in the personalised interactive system because user actions are not traceable.
3. Existing user models cannot address some critical challenges when used in an interactive task. E.g. cognitive model will have scalability challenge, empirical model will have user knowledge retrieval challenge, and usage model will have usage information interpretation challenge.
4. Personalisation of live sports events viewing system can be achieved only when the challenges are addressed in various perspectives. The interaction should be redesigned to support the personalisation requirements, e.g. existing zoomable user interface is not able to support video content and a video zoomable

interface should be proposed. User preference retrieval is another challenge as user preference is changing during live events. And the preference adaptation should also help reduce the interaction complexity rather than merely offering an adaptive system reaction.

5. An evaluation of personalisation is critical to ensure the performance of personalisation. As of today, there is neither clear standard of personalisation quality nor evaluation framework.

Chapter 3

3 User Interaction in a Next Generation A-V Player

User requirements play a dominant role in determining the interactive tasks needed within an application domain such as the use of a next generation A-V player for live sports events. User requirements were collected as part of the FP7 EU Project My eDirector 2012 survey and analysed. First, these user requirements are used to define the core user task models. Then, in order to address the challenge of modelling user task that support personalisation, improved user task models are proposed that enables tasks to be personalised.

3.1 Interaction Requirements

Requirements are crucial to any system design and modelling. Requirements do not necessarily describe what and how reality is, rather, they model reality as it should be (Dzida, 1998). The increasing transmission rates available to access video content coupled with the use of advanced rich Internet application technologies enables service providers including television broadcasters to offer viewers richer content while giving them more flexible means to interact with the content. Hence, Web based video broadcasting platforms are becoming much more popular, e.g. BBC's iPlayer, NBC's Sunday Night Football Extra Player etc. The My-e-Director 2012 EU FP7 project survey of user requirements (My-e-Director 2012, 2008) conducted by commercial partners such as ATOS and BBC, in which 445 users from 7 countries (UK, Greece, Spain, Portugal, France, Germany and Italy) participated, reveals that a sizable majority of users use Internet ready terminals such as PCs and laptops. The original questions from My-e-Director 2012 user Requirements Deliverable (2008) are attached in Appendix C. The survey also highlighted four interaction requirements for the next generation A-V player: sports events selection, multi-angle viewing of events, selective target zooming and time-shift viewing of events (which refers to recording an event and to watching this at a more convenient time).

Viewers also tend to have varying preferences for the types of sports they prefer to watch. According to a BBC research report (BBC, 2004) of how the UK and other nations viewed a previous Olympics, the Athens Olympics, viewers are particularly interested in certain sports, with athletics appearing, by far, as the most popular Olympic sport. Other sports of particular interest vary significantly between nations. In the My-e-

Director 2012 survey (2008), participants from 7 European countries expressed their interests for sports events in general. The three most popular sports were football (58%), basketball (49%), tennis and athletics (38%).

In the user requirements survey (My-e-Director 2012, 2008), 62% of participants wanted support for camera switching enabling them to view events from different angles. The most popular views in order of preference were: an aerial view of the event (~50%), close-ups of specific athletes, and views of front runners and views of the main track. Note also for viewers at a live event (on-site), rather than those that view it remotely (off-site), the ability to switch camera with mobile phones is the most desired interaction (45%).

Time-shift viewing includes replay, slow motion, rewind, fast forward etc. In the survey report (My-e-Director 2012, 2008), a replay is preferred by 65% of the users which is the most supported interaction among others. 47% of the participants are willing to pay to have the ability to replay live content. 41% of the on-site participants would like to be able to replay significant moments in slow motion, making this the second most favoured feature after multi-angle viewing for on-site viewers.

In addition to viewing a greater variety of coverage, viewers are also interested in viewing specific targets such as athletes in greater detail. There are different approaches to achieve this selective target zooming. In the survey (My-e-Director 2012, 2008), two questions relate to this, one concerns zooming and the other concerns tagging and tracking athletes. Tagging and tracking are supported by 36% of the users. 45% of users expressed a preference for zooming – making this the third most preferred feature. Although it is not clear why users prefer zooming to tagging and tracking athletes, it may be that users are more familiar with zooming.

To sum up the survey results, sports events selection, multi-angle viewing of events, selective target zooming and time-shift viewing are the four most advocated and expected interactive tasks of the next generation A-V player for live sports events viewing. In the next section, a new interactive task model is proposed in order to model these interactions and to identify the sub tasks that require personalisation.

3.2 Interactive User Task Model

User task modelling aims to design a model that can assist users to execute a task more effectively. The advanced user tasks discussed in the previous section can involve different sub tasks and some physical operation demanding sub tasks that need further personalisation to reduce repetitive operations. Hence, a user task model is needed that is able to:

1. explicitly describe the task execution constraints between sub tasks,
2. be described by a formal logic so that the model can be validated before implementation,
3. identify the critical sub tasks for personalisation.

For the first requirement, existing task models such as HTA, GOMS and CTTE all support a description of execution constraints between sub tasks (see section 2.2.2). However, both HTA and GOMS lack a standard annotation whereas CTTE defines a set of standard graphical annotations. For the other two requirements, i.e. 2 and 3, few existing models are capable of achieving them. Table 3-1 evaluates the existing mainstream task models (see section 2.2.2) with respect to these three requirements.

Table 3-1 Task Models Evaluation Matrix

Task Model	**Task Execution Constraints Representation**	**Formal Logic Representation**	**Personalisation Identification Support**
HTA	Simple 'if...else...' rules	X	X
GOMS	natural language	X	X
User-task Elicitation Tool	X	X	X
CTTE	graphical notations	X	X

The proposed interactive user task model in this thesis leverages CTTE's graphical representation to describe the task execution constraints between sub tasks. It also supports a formal logic representation of the task description using computational tree logic. At last, the critical personalisation sub tasks can be identified. In the next few sections, the proposed task model is detailed.

3.2.1 Task Description

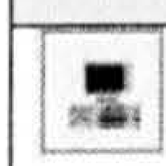
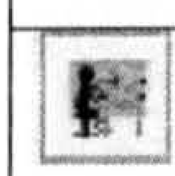
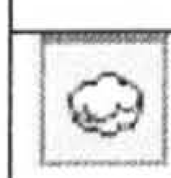
A task model defines the syntax or structure of the task, the types of operators (i.e. relational symbols between sub tasks) that act on tasks and the semantic meaning of theses operators. The syntax of task model is either an informal natural language, e.g. GOMS, or it is a graphical notation, e.g. CTTE or a formal notation.

The proposed interactive task model here leverages both CTTE graphical annotations and computational tree logic (CTL) formal notations to describe the task. The former is mainly used to define the type of tasks and to constrain the execution between sub tasks and the latter is mainly used to allow the task model to be validated.

3.2.1.1 Graphical Representation

A graphical structure is presented to describe how user tasks are arranged and to define the actors for these tasks. The meaning of the symbols and operators is listed in Table 3-2.

Table 3-2 Task notations and relations notations

Symbols and Operators	Semantic Meaning
	**System Task** involves only system actions
	**Interactive task** involves both user actions and system actions
	**Abstract task** involves both system tasks and interactive tasks
Enable: A >> B	Task B is not able to start unless Task A is performed
Choice: A[] B	Task A and B are both available but only one is allowed to be performed at a time
Enable with information passing: A []>> B	Task A must complete and pass information to B at B's start

A simple example for "sports events selection" using the hierarchical structure task representation is illustrated in Figure 3-1.

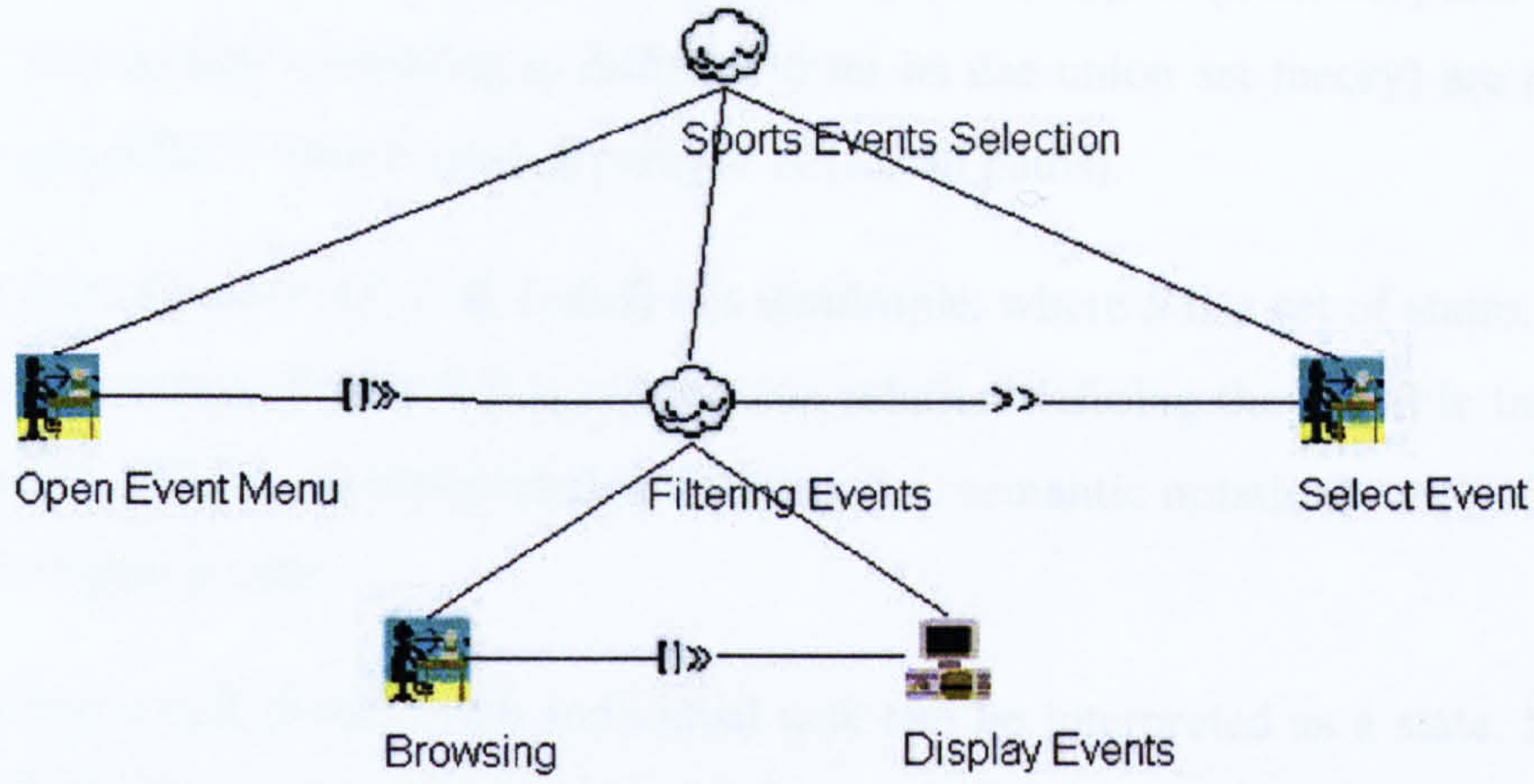


Figure 3-1 Graphical representation of a sports events selection task model

This consists of four sub tasks namely 'open event menu', 'browsing', 'display events' and 'select events' organised in two sub-levels. 'Browsing' and the 'display events' are both second level sub tasks which belong to the first level sub task 'filtering events'. When a task involves both interactive and system sub tasks, it is called an abstract task such as 'Sports Events Selection' and 'Filtering Events'. When a task is solely performed by the system such as 'Display Events', it is called a system task. When a non-abstract task involves user interactions, it is called an interactive task.

3.2.1.2 Formal Logic Representation

Task relations can also be expressed using a formal logic. The interactive task model here draws upon the concept of Computational Tree Logic or CTL (Clarke and Emerson, 1982) to express the task transition processes. CTL's semantics are interpreted with reference to Kripke Structures (Kripke, 1971). This is a transition system that describes the state transition behaviour. The syntax of CTL defines transitions according to the following rules:

1) p is state-formula, which defines the state
2) If Φ is a state-formula, then $\neg\, \Phi$ is a state-formula
3) If $\Phi \wedge \Psi$ is a state-formula, then $\Phi \vee \Psi$ is a state-formula
4) If φ is a path-formula, then $E\varphi$ and $A\varphi$ are state-formulas
5) If Φ is a state-formula, then $X\ \Phi$ or $F\ \Phi$ or $G\ \Phi$ or $U\ \Phi$ is a path-formula
6) If Φ and Ψ is a state-formula, then $\Phi \cup \Psi$ are a state-formulas

Linear temporal operators X (next state), F (in a future state), G (Globally) and U (until, note that this symbol's meaning is different from its use union set theory) are preceded by a path quantifier E (there exists a path) or A (for all paths).

A Kripke Structure $M = (S, I, R, Label)$ is a quadruple, where S is a set of states, $I \subseteq S$ is a set of initial states, $R \subseteq S\ X\ S$ is a transition relation defining the possible transitions among states. *Label* is an interpretation function (i.e. semantic notation) on S, where $p \in Label(s)$ is called p state.

In the proposed task model, each individual task can be interpreted as a state. For each transition between tasks, it starts with an initial task and reaches a final task.

The state-formulas can be defined by atomic propositions:

$AP = \{P_0\text{= parent}, P_1\text{=abstract}, P_2\text{=interaction}, P_3\text{= system}\}$

Figure 3-2 shows the Kripke Structure of Sports Events Selection graphically. Note that this figure is used to facilitate the understating of the Kripke Structure only. The structure itself can also be expressed formally.

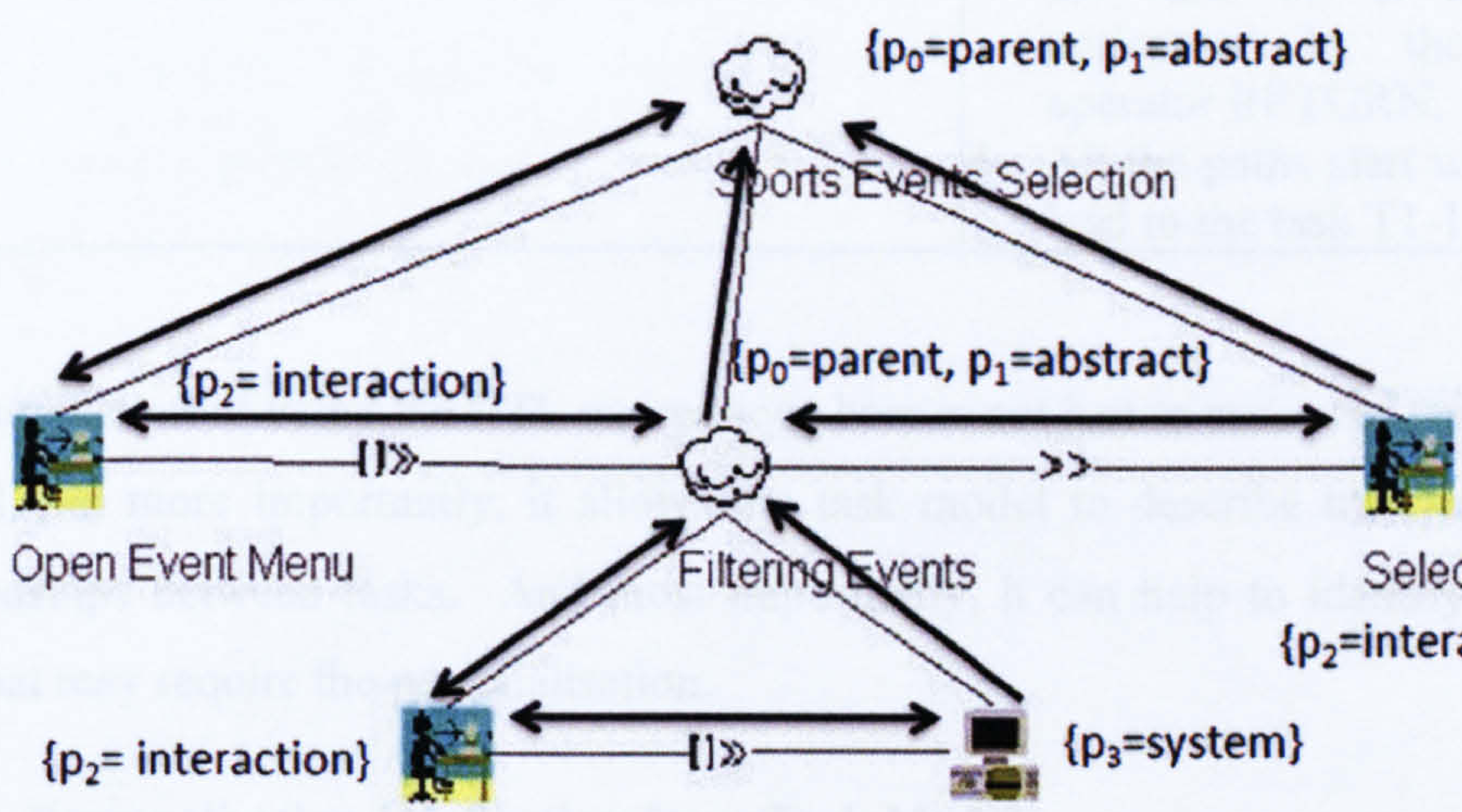


Figure 3-2 Event selection task in Kripke structure

In Figure 3-2, a hidden transition operator 'RETURN' is introduced, which is represented as an arrow pointing to the task itself. With this operator, each child task can be aborted and return to its parent task or to a previous task. To illustrate how to express the state transition with CTL, two Abstract user tasks 'Sports Events Selection' (T0) and 'Filtering Events' (T1) are taken as the examples as shown in Table 3-3. T0-1 denotes the 'Open Event Menu', T0-2 denotes 'Select Event', T1-1 denotes 'Browsing' and T1-2 denotes 'Display Events'. Here, in the task notation 'TX-Y', X denotes the Abstract task and Y denotes the sub task belongs to X.

Table 3-3 Task Relations as CTL Expressions

CTL Expression	**Meaning**
T0 as Initial Task: *AG(task = T0) ⇒ E (U(task = T1))*	All the paths starting with Task T0 lead to Task T1
T1 as Initial Task: *AG(task = T1) ⇒ E((transition = >>) U (task = T0-2))* *And* *AG(task = T1) ⇒ E((transition = RETURN) U (task = T0-1))* *And* *AG(task = T1) ⇒ E((transition = RETURN) U (task = T0))* *And* *AG(task = T1) ⇒ E (U(task = T1-1))*	Task T1 has four paths to go to T0-2, returns to T0-1, T0 or leads to T1-1. 1. All the paths start with T1 until the task T0-2 is reached are performed by the transition operator >> (Enable). 2. All the paths start with T1 until the task T0-1 is reached are performed by the transition operator RETURN. 3. All the paths start with T1 until the task T0 is reached are performed by the transition operator RETURN. 4. All the paths start with Task T1 lead to the task T1-1

The implication of using the CTL expressions here is not just to make the task modelling formal, but more importantly, it allows the task model to describe traceable temporal relationships between tasks. And most importantly, it can help to identify the critical task that may require the personalisation.

3.2.2 Personalisation Implication from Task Model

User task models in general do not explicitly support personalisation. One of the important features of the interactive task model proposed here is it is able to indicate the critical sub task requires personalisation.

In order to find out the critical sub task that requires personalisation within an abstract task, three steps are required. The algorithm is shown in Table 3-4. The general idea of this algorithm is to analyse the task model complexity in terms of sub task levels and information entropy of each concerned task where the information entropy is defined as the information load that task contains which eventually depends on the number of paths (i.e. links from/to other tasks) that go through the task. Task levels can be identified from the graphical representation of the task. Information entropy can be obtained from the CTL representation of the task.

Table 3-4 Algorithm to identify critical sub task

Step1: Identify the number of sub task levels $L = \{L0, \ldots, Ln\}$, where top level = $L0$ Identify the interactive task $Ti = \{Ti0, \ldots, Tin\}$ Identify the Abstract task $Tc = \{Tc0, \ldots, Tcn\}$ Identify the paths for overall and interactive task i.e. $Pin = \{ P0, \ldots, Pn \}$, $Pcn = \{ P0, \ldots, Pn \}$
Step2: Calculate the information entropy of each task $H(T_n) = -\sum_{i=1}^{n} p(P_n) \log p(P_n)$
Step3: Compare the information entropy among tasks at the same task level LOOP $T_{ps} = Max(H(T_0), \ldots, H(T_n) \subseteq L_j), j++ where, j < L$ END LOOP where j=L Critical sub task will be T_{ps} given T_{ps} is an interactive task

In Figure 3-3, each task node is labelled with possible CTL paths to that node. Due to the fact that the task model describes the initial status of each task, there will be an equal chance for each path to go to its associated task node and hence the possibility of each transition path will be $1/N$ where N is the number of paths associated to each task node.

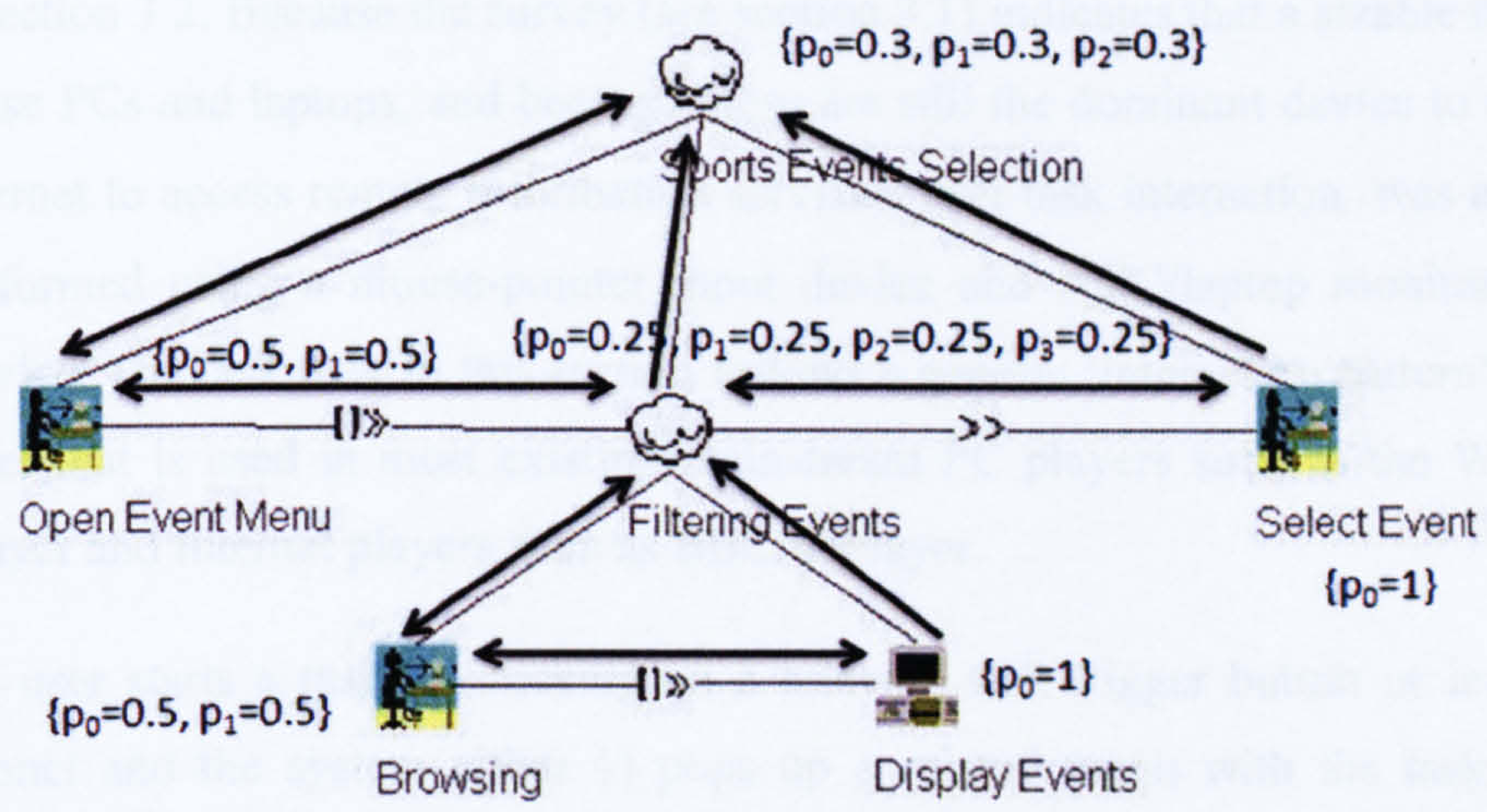


Figure 3-3 Task model annotated with CTL transition path possibility

At the top level, only one node is available that is 'Sports Events Selection'. It has an information entropy value of 0.47. Because it is the only node at the top level it has an information entropy value greater than zero. The lower level of sub tasks will be examined in a similar way. Figure 3-4 illustrates how the critical task that requires personalisation can be found. The search route starts from the 'Sports Events Selection' and moves to 'Filtering Events' and lastly leads to 'Browsing'.

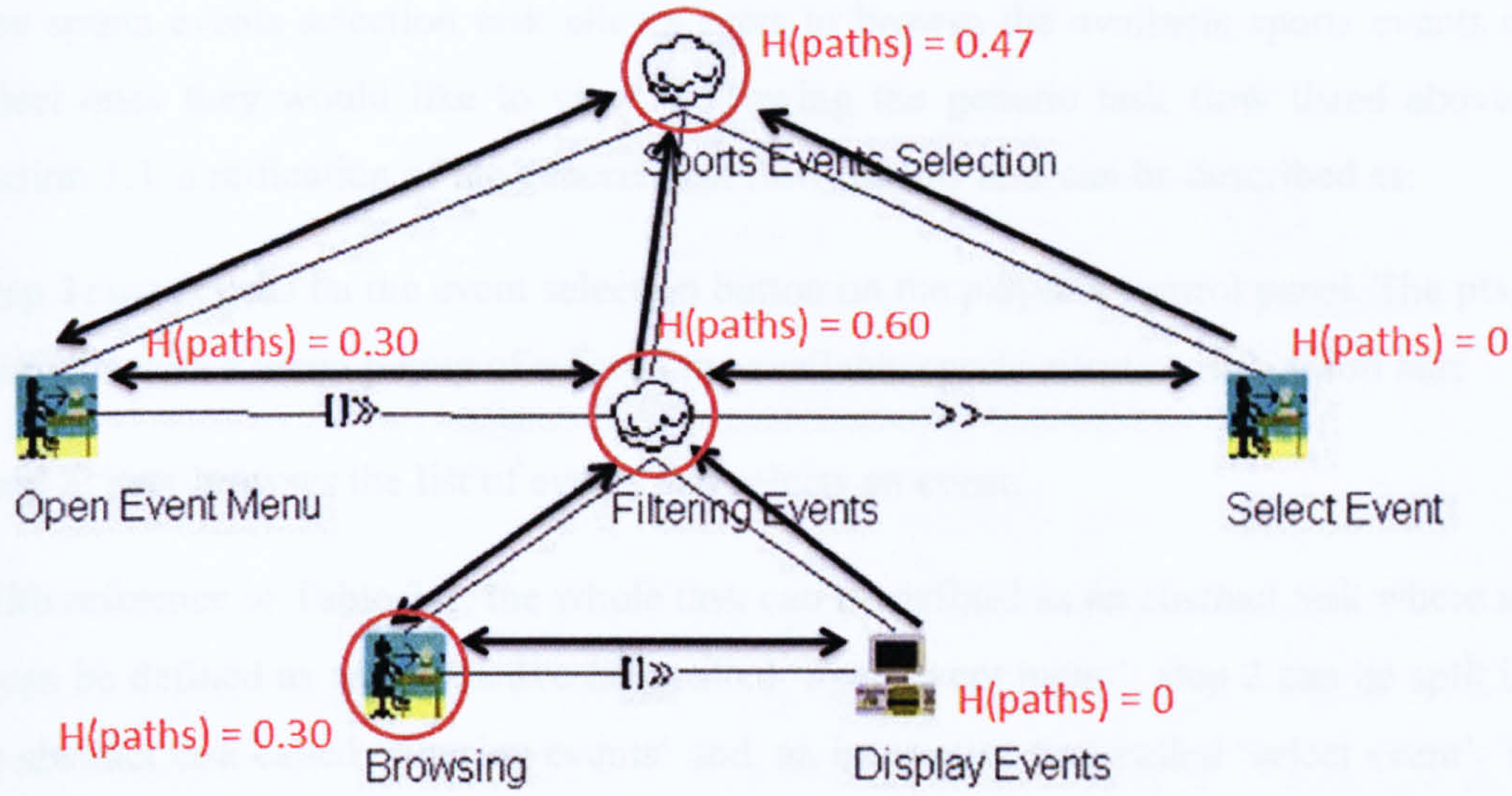


Figure 3-4 Task nodes with information entropy values for a 'Browsing' task that requires personalisation

3.3 Task Model Application

Here the user tasks identified in section 3.1 are modelled using the proposed task model given in section 3.2. Because the survey (see section 3.1) indicates that a sizable majority of users use PCs and laptops, and because these are still the dominant device to connect to the Internet to access remote information services, user task interaction was assumed to be performed using a mouse-pointer input device and a PC/laptop monitor as the output device. The task flow in this section follows a generic 'interaction pattern' for an A-V player that is used in most existing mainstream PC players such as the Windows Media Player and Internet players such as BBC's iPlayer.

Step 1: a user starts a task by clicking on a relevant task trigger button or icon on a control panel and the system either 1) pops up a related menu with the task related options shown or 2) it directly triggers the associated system task functions;

Step 2: if a task related menu is presented, user enters some task-specific options and system triggers an associated system task function

This task flow is a common interaction pattern to trigger system tasks and can be easily learned. The learnability amongst other aspects is further evaluated with a usability test presented in section 6.3 to justify this task flow.

3.3.1 Sports Events Selection Task

The sports events selection task allows users to browse the available sports events and select ones they would like to view. Following the generic task flow listed above in section 3.3, a reification of the generic task flow for this task can be described as:

Step 1: user clicks on the event selection button on the player's control panel. The player responds with a menu popup of a list of the available sports events with a scroll bar;

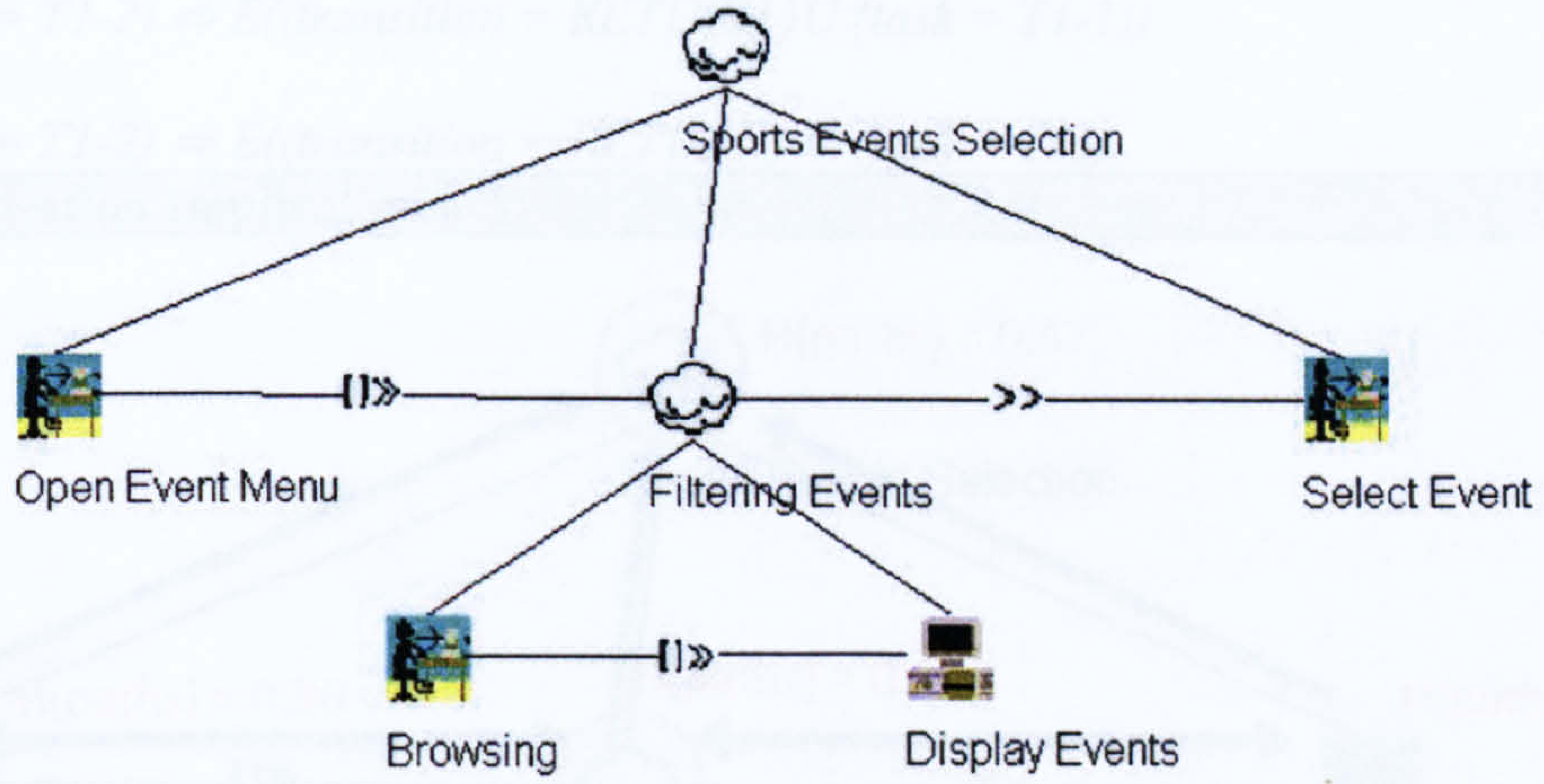
Step 2: user browses the list of events and selects an event.

With reference to Table 3-2, the whole task can be defined as an abstract task where step 1 can be defined as an interactive task called 'open event menu'; step 2 can be split into an abstract task called 'filtering events' and an interactive task called 'select event'. The 'filtering events' can further consist of one interactive task called 'browsing' and one system task called 'display events'. Thus, a sports events selection task can be modelled as shown in Table 3-5 in which the sub tasks are expressed both in CTTE diagrams and CTL language and the sub task 'Browsing' is identified as the critical sub task that requires personalisation. An alternative task model application here is to view step 2 as an abstract task that can consist of three sub tasks i.e. 'browsing', 'display events' and 'select event'. In this case, the identified critical personalisation sub task will be the same to the original application, i.e. 'browsing'.

Table 3-5 Sports events selection task model

Task Descriptions

Graphical Representation



Task 0 Sports Events Selection: an abstract task that represents the whole task

Task 0-1 Open Event Menu: an interactive task that allows user to launch an event menu

Task 1 Filtering Events: an abstract task that follows Task 0-1, the information passed from Task 0-1 includes the time that determines which live events are currently being offered.

Task 1-1 Browsing events: an interactive task that allows the user to visually go over the events list.

Task 1-2 Display Events: a system task that displays the events when a user browses the event list.

Task 0-2 Select event from displayed events: this interactive task allows the user to finalise an event selection.

Textual Representation

T0 as Initial Task:

$AG(task = T0) \Rightarrow E\ (U(task = T1))$

T0-1 as Initial Task:

$AG(task = T0\text{-}1) \Rightarrow E((transition = []>>)\ U\ (task = T1))$

AND

$AG(task = T0\text{-}1) \Rightarrow E((transition = RETURN)\ U\ (task = T0))$

T0-2 as Initial Task:

$AG(task = T0\text{-}2) \Rightarrow E((transition = RETURN\ )U\ (task = T1))$

AND

$AG(task = T0\text{-}2) \Rightarrow E((transition = RETURN)\ U\ (task = T0))$

T1 as Initial Task:

$AG(task = T1) \Rightarrow E((transition = >>)\ U\ (task = T0\text{-}2))$

AND

$AG(task = T1) \Rightarrow E((transition = RETURN)\ U\ (task = T0\text{-}1))$

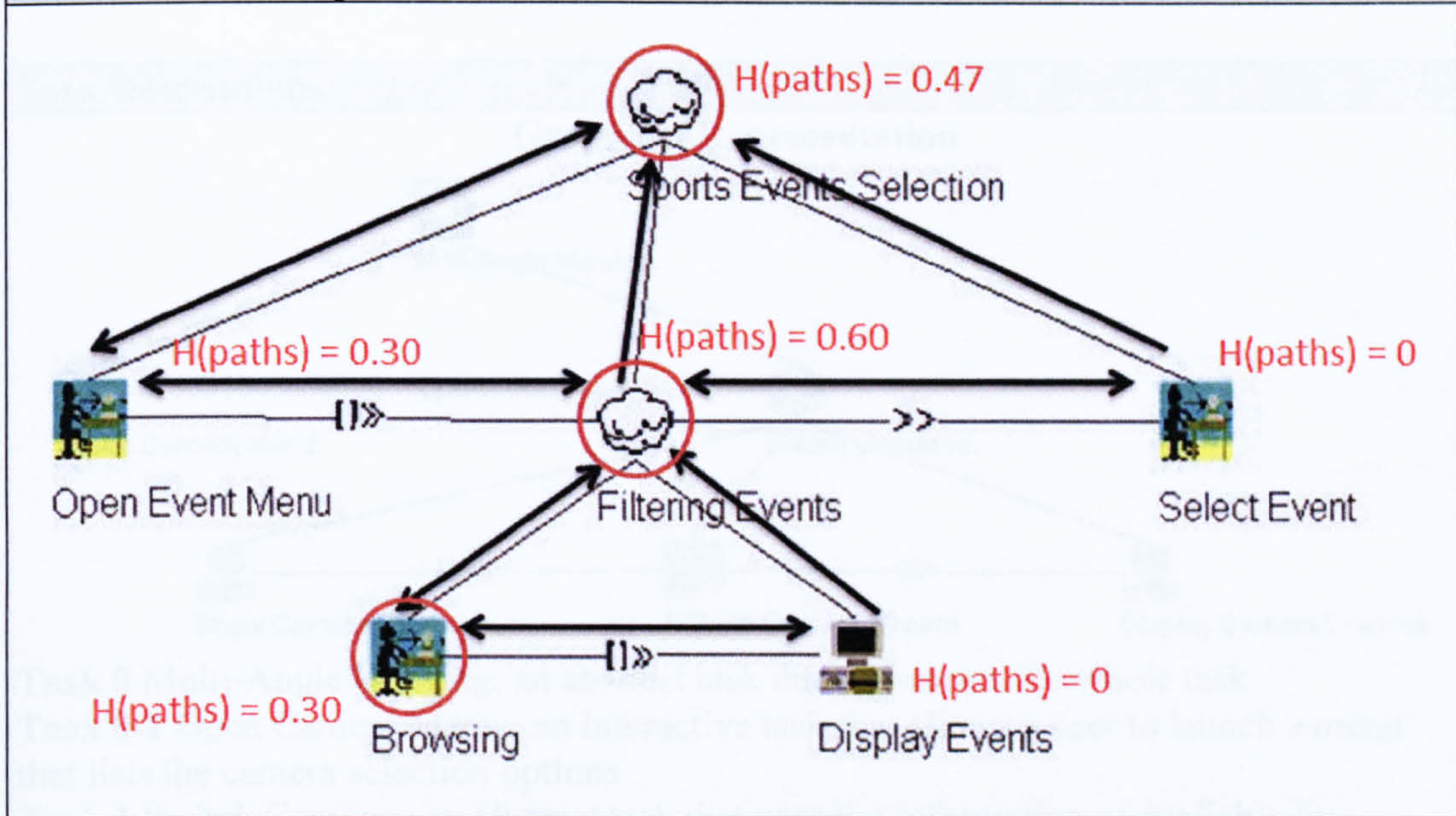
AND

$AG(task = T1) \Rightarrow E((transition = RETURN)\ U\ (task = T0))$

AND

$AG(task = T1) \Rightarrow E\ (U(task = T1\text{-}1))$

T1-1 as Initial Task:
$AG(task = T1\text{-}1) \Rightarrow E((transition = RETURN)\ U\ (task = T1))$
AND
$AG(task = T1\text{-}1) \Rightarrow E((transition = []>>)\ U\ (task = T1\text{-}2))$
T1-2 as Initial Task:
$AG(task = T1\text{-}2) \Rightarrow E((transition = RETURN)\ U\ (task = T1\text{-}1))$
AND
$AG(task = T1\text{-}2) \Rightarrow E((transition = RETURN)\ U\ (task = T1))$
Personalisation Implication



3.3.2 Multi-Angle Viewing Task

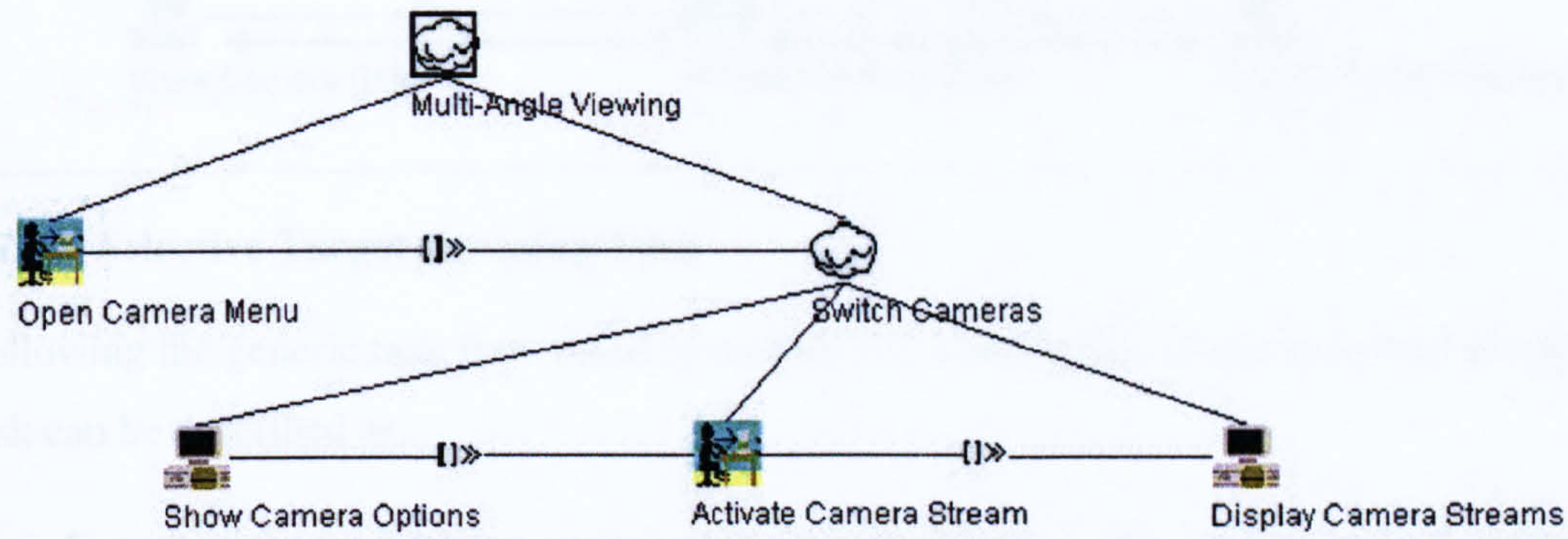
Following the generic task flow listed in section 3.3, a reification of this task flow for this task can be described as:

Step 1: user clicks on a button on the player's control panel and player responds with a panel;

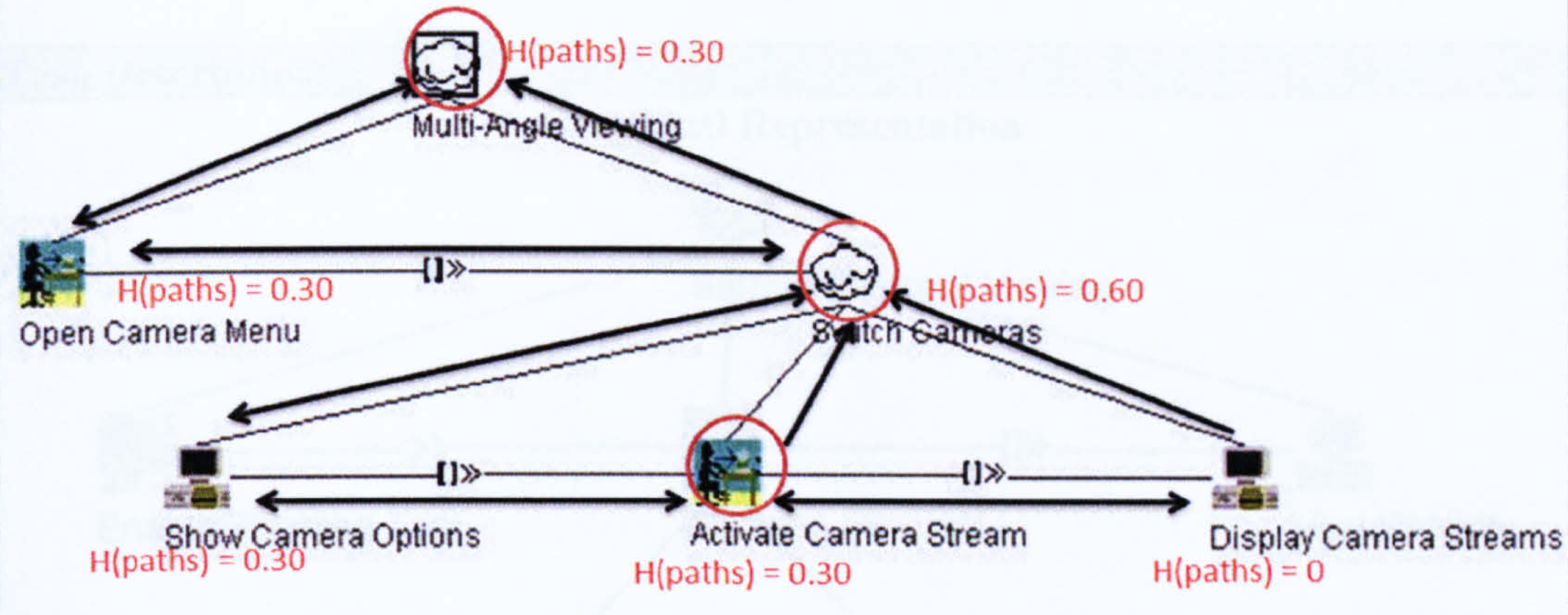
Step 2: based upon the available camera options in the panel, user switches around the cameras and player displays the switched cameras.

According to Table 3-2, the whole task can be defined as an abstract task where step 1 can be defined as an interactive task; step 2 can be defined as an abstract task. Here, step 1 is named as 'open camera menu', step 2 consists of one system task called 'show camera options', one interactive task called 'activate camera stream' and a system task named 'display camera streams'. There are two alternative task model applications here, i.e. one is to combine view step 2 as an abstract task which consists of three sub tasks and the other is to split step 2 into one abstract task (i.e. 'show camera options' and 'activate camera stream') and one system task 'display camera streams'. The critical personalisation sub task in both cases will be the same, i.e. 'activate camera stream'. Table 3-6 shows the multi-angle viewing task model that views step 2 as an abstract task in which the sub tasks are expressed both as CTTE diagrams and in the CTL language. The sub task 'Activate camera stream' is identified as the critical sub task requires personalisation.

Table 3-6 Multi-angle viewing task model

Task Descriptions
Graphical Representation

Task 0 Multi-Angle Viewing: an abstract task that represents the whole task **Task 0-1** Open Camera Menu: an interactive task that allows a user to launch a menu that lists the camera selection options **Task 1** Switch Cameras: an abstract task that uses the information of available live cameras passed from the Task 0-1. **Task 1-1** Show Camera Options: a system task that presents a GUI of the available cameras. **Task 1-2** Activate Camera Stream: an interactive task that allows a user to start a selected camera stream. **Task 1-3** Display Camera Stream: a system task that displays the video content based upon the selected camera stream.
Textual Representation
T0 as Initial Task: $AG(task = T0) \Rightarrow E\ (U(task = T0\text{-}1))$ ***T0-1 as Initial Task:*** $AG(task = T0\text{-}1) \Rightarrow E((transition = ()>>)\ U\ (task = T1))$ $AND\ AG(task = T0\text{-}1) \Rightarrow E((transition = RETURN)\ U\ (task = T0))$ ***T1 as Initial Task:*** $AG(task = T1) \Rightarrow E(U\ (task = T1\text{-}1))$ $AND\ AG(task = T1) \Rightarrow E((transition = RETURN)\ U\ (task = T0\text{-}1))$ $AND\ AG(task = T1) \Rightarrow E((transition = RETURN)\ U\ (task = T0))$ ***T1-1 as Initial Task:*** $AG(task = T1\text{-}1) \Rightarrow E((transition = RETURN\ )U\ (task = T1))$ $AND\ AG(task = T1\text{-}1) \Rightarrow E((transition = ()>>)\ U\ (task = T1\text{-}2))$ ***T1-2 as Initial Task:*** $AG(task = T1\text{-}2) \Rightarrow E((transition = RETURN\ )U\ (task = T1\text{-}1))$ $AND\ AG(task = T1\text{-}2) \Rightarrow E((transition = RETURN)\ U\ (task = T1))$ $AND\ AG(task = T1\text{-}2) \Rightarrow E((transition = ()>>)\ U\ (task = T1\text{-}3))$ ***T1-3 as Initial Task:***

$AG(task = T1\text{-}3) \Rightarrow E((transition = RETURN)\,U\,(task = T1\text{-}2))$ $AND\ AG(task = T1\text{-}3) \Rightarrow E((transition = RETURN)\,U\,(task = T1)$
Personalisation Implication



3.3.3 Selective Target Zooming Task

Following the generic task flow listed in section 3.3, a reification of this task flow for this task can be described as:

Step 1: user clicks on a button on the player's control panel. The player responds with a message box showing a task instruction on how to zoom-in or zoom-out, e.g. double-click on left mouse button on screen to zoom-in and single-click on left mouse button to zoom-out;

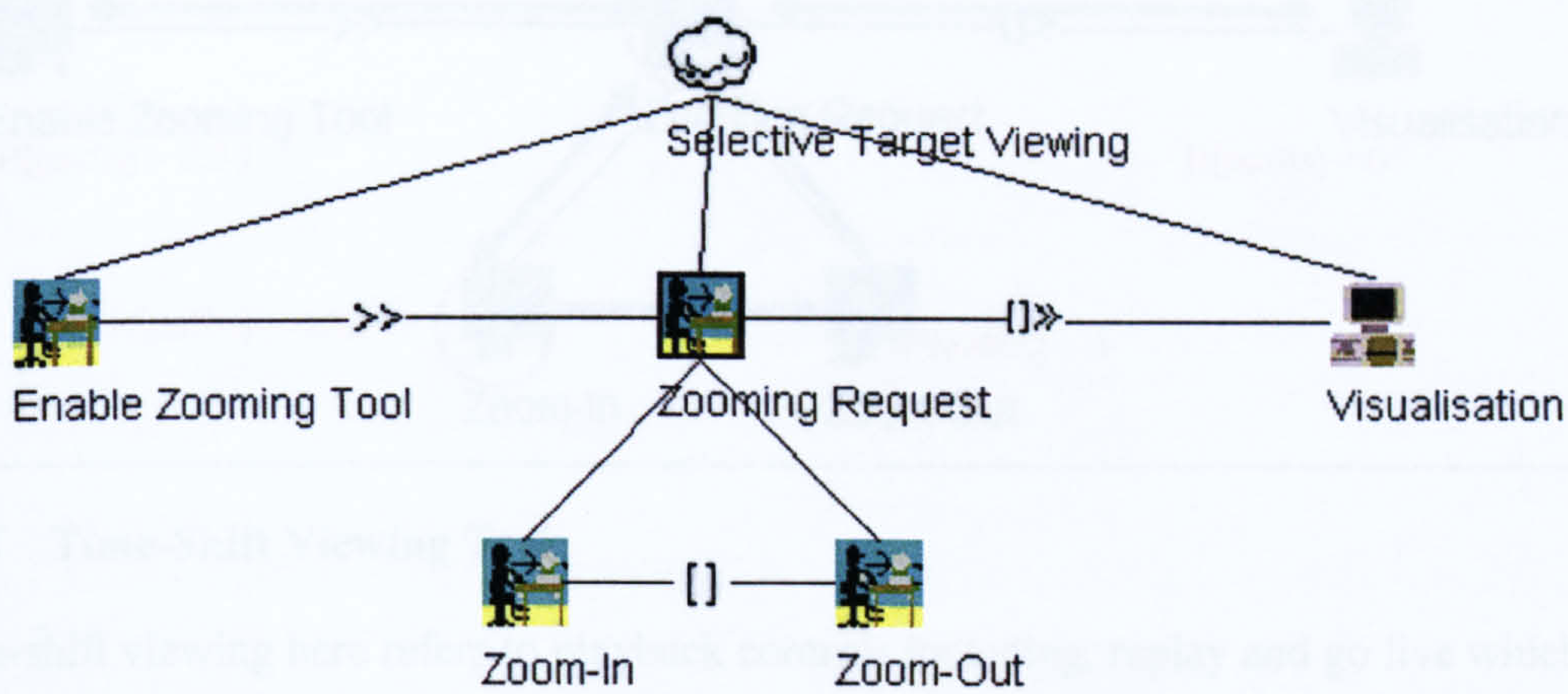
Step 2: user clicks mouse button on screen to zoom-in or zoom-out;

According to Table 3-2, the whole task can be defined as an abstract task where step 1 can be defined as an interactive task named as 'enabling zooming tool'; step 2 can be split into two interactive tasks called 'zoom-in' and 'zoom-out' and a system task called 'visualisation'. Given the keyboard and mouse are the main input devices to the player, this design aims to contextualise the use of the mouse cursor so that "clicks" on the screen cannot lead to other functions such as full screen or live pause. Table 3-7 shows the model of the zooming task in which the sub tasks are expressed both in CTTE diagrams and CTL language and the sub task 'Zoom in' is identified as the critical sub task that requires personalisation.

Table 3-7 Selective target zooming task model

Task Descriptions

Graphical Representation



Task 0 Selective target zooming: an abstract task that represents the whole task
Task 0-1 Enable Zooming Tool: an interactive task that allows a user to activate the zooming tool
Task 1 Zooming Request: an interactive task that follows the Task 0-1. This allows a user to use the zooming tool, i.e. either zoom in or zoom out.
Task 1-1 Zoom-in: an interactive task that allows a user to amplify the target of interest.
Task 1-2 Zoom-out: an interactive task allows a user to reduce the magnification level back to the pre zoom-in state.
Task 0-2 Visualisation: a system task that displays a video stream based upon the user zooming request information, i.e. zoom-in or zoom-out.

Textual Representation

T0 as Initial Task:
$AG(task = T0) \Rightarrow E\,(U(task = T1))$

T0-1 as Initial Task:
$AG(task = T0\text{-}1) \Rightarrow E((transition = >>)\ U\ (task = T1))$
$AND\ AG(task = T0\text{-}1) \Rightarrow E((transition = RETURN)\ U\ (task = T0))$

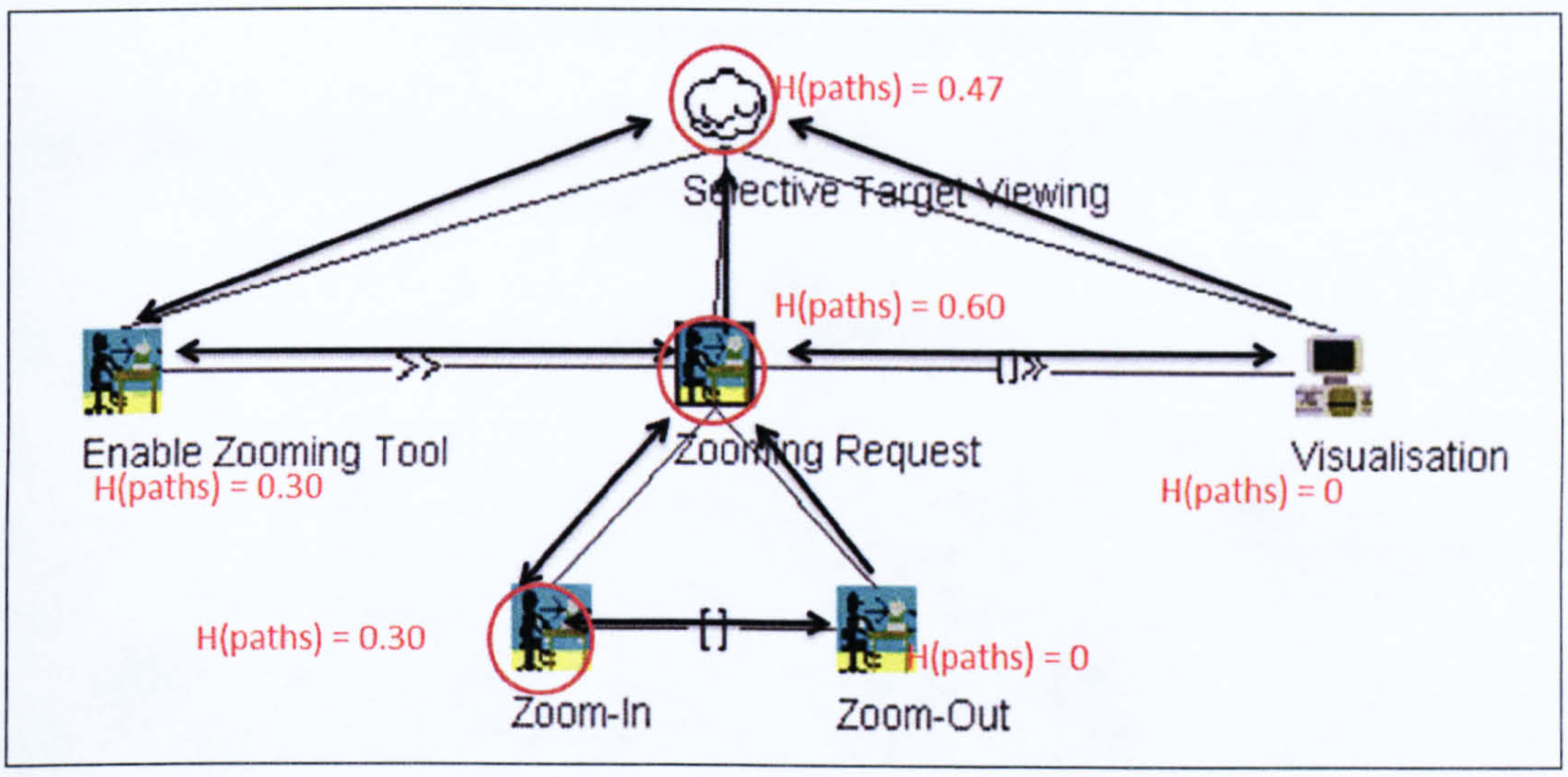
T0-2 as Initial Task:
$AG(task = T0\text{-}2) \Rightarrow E((transition = RETURN\,)U\ (task = T1))$
$AND\ AG(task = T0\text{-}2) \Rightarrow E((transition = RETURN)\ U\ (task = T0))$

T1 as Initial Task:
$AG(task = T1) \Rightarrow E((transition = ()>>)\ U\ (task = T0\text{-}2))$
$AND\ AG(task = T1) \Rightarrow E((transition = RETURN)\ U\ (task = T0\text{-}1))$
$AND\ AG(task = T1) \Rightarrow E((transition = RETURN)\ U\ (task = T0))$
$AND\ AG(task = T1) \Rightarrow E\,(U(task = T1\text{-}1))$

T1-1 as Initial Task:
$AG(task = T1\text{-}1) \Rightarrow E((transition = RETURN\,)U\ (task = T1))$
$AND\ AG(task = T1\text{-}1) \Rightarrow E((transition = ())\ U\ (task = T1\text{-}2))$

T1-2 as Initial Task:
$AG(task = T1\text{-}2) \Rightarrow E((transition = ()\,)U\ (task = T1\text{-}1))$
$AND\ AG(task = T1\text{-}2) \Rightarrow E((transition = RETURN)\ U\ (task = T1))$

Personalisation Implication



3.3.4 Time-Shift Viewing Task

Time-shift viewing here refers to playback controls including, replay and go live which is most popular feature as discussed in section 3.1. Following the generic task flow listed in section 3.3, a reification of this task flow for this task can be described as:

Step 1: a user clicks on a button on the player's control panel. The player pops up a menu which allows a user to capture the event highlights;

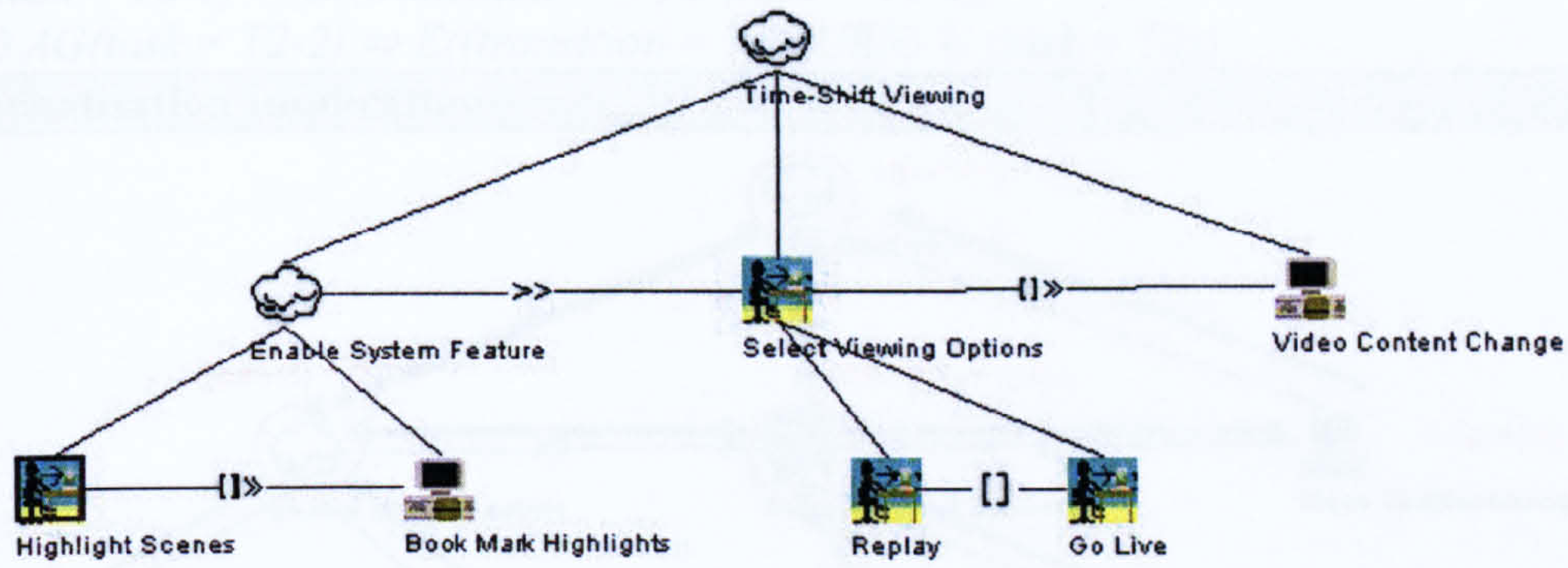
Step 2: a user chooses to replay the highlights or to restore the view to the live streams view. The player streams the highlights or the live content respectively.

With reference to Table 3-2, the whole task can be defined as an abstract task. Step 1 can be defined as an abstract task which consists of an interactive task called 'highlight scenes' and one system task called 'bookmark highlights'; step 2 can be split into two interactive tasks called 'replay' and 'go live' and a system task called 'video content change'. The task model is shown in Table 3-8 in which the sub tasks are expressed both as CTTE diagrams and using the CTL language. Two sub tasks 'Highlight Scenes' and 'Replay' are identified as the critical sub tasks that require personalisation.

Table 3-8 Time-shift viewing task model

Task Descriptions

Graphical Representation



Task 0 Time-shift Viewing: an abstract task that represents the whole task
Task 1 Enable System Feature: an interactive task that allows a user to highlight the current live stream for later time-shift viewing.
Task 1-1 Highlight Scenes: an interactive task that allows a user to highlight scenes during a live event.
Task 1-2 Bookmark Highlights: a system task that bookmarks the user highlighted scenes.
Task 2 Select Viewing Options: an interactive task that follows Task 0-1 and allows a user to start either replay or to go back to live.
Task 2-1 Replay: an interactive task that allows a user to replay the selected highlights of a current live event.
Task 2-2 Go Live: an interactive task allows a user to catch the live scenes of a current event.
Task 0-1 Video Content Change: a system task that render the proper images of current live stream based upon the user selected viewing option, i.e. replay or go live

Textual Representation

T0 as Initial Task:
AG(task = T0) ⇒ E (U(task = T1))
T1 as Initial Task:
AG(task = T1) ⇒ E((transition = >>) U (task = T2))
AND AG(task = T1) ⇒ E((transition = RETURN) U (task = T0))
AND AG(task = T1) ⇒ E (U(task = T1-1))
T1-1 as Initial Task:
AG(task = T1-1) ⇒ E((transition = ()>>) U (task = T1-2))
AND AG(task = T1-1) ⇒ E((transition = RETURN) U (task = T1))
T1-2 as Initial Task:
AG(task = T1-2) ⇒ E((transition = RETURN) U (task = T1-1))
AND AG(task = T1-2) ⇒ E((transition = RETURN) U (task = T1))
T0-1 as Initial Task:
AG(task = T0-1) ⇒ E((transition = RETURN)U (task = T2))
AND AG(task = T0-1) ⇒ E((transition = RETURN) U (task = T0))
T2 as Initial Task:
AG(task = T2) ⇒ E((transition = ()>>) U (task = T0-1))
AND AG(task = T2) ⇒ E((transition = RETURN) U (task = T1))
AND AG(task = T2) ⇒ E((transition = RETURN) U (task = T0))
AND AG(task = T2) ⇒ E (U(task = T2-1))

T2-1 as Initial Task:

$AG(task = T2\text{-}1) \Rightarrow E((transition = RETURN)\,U\,(task = T2))$

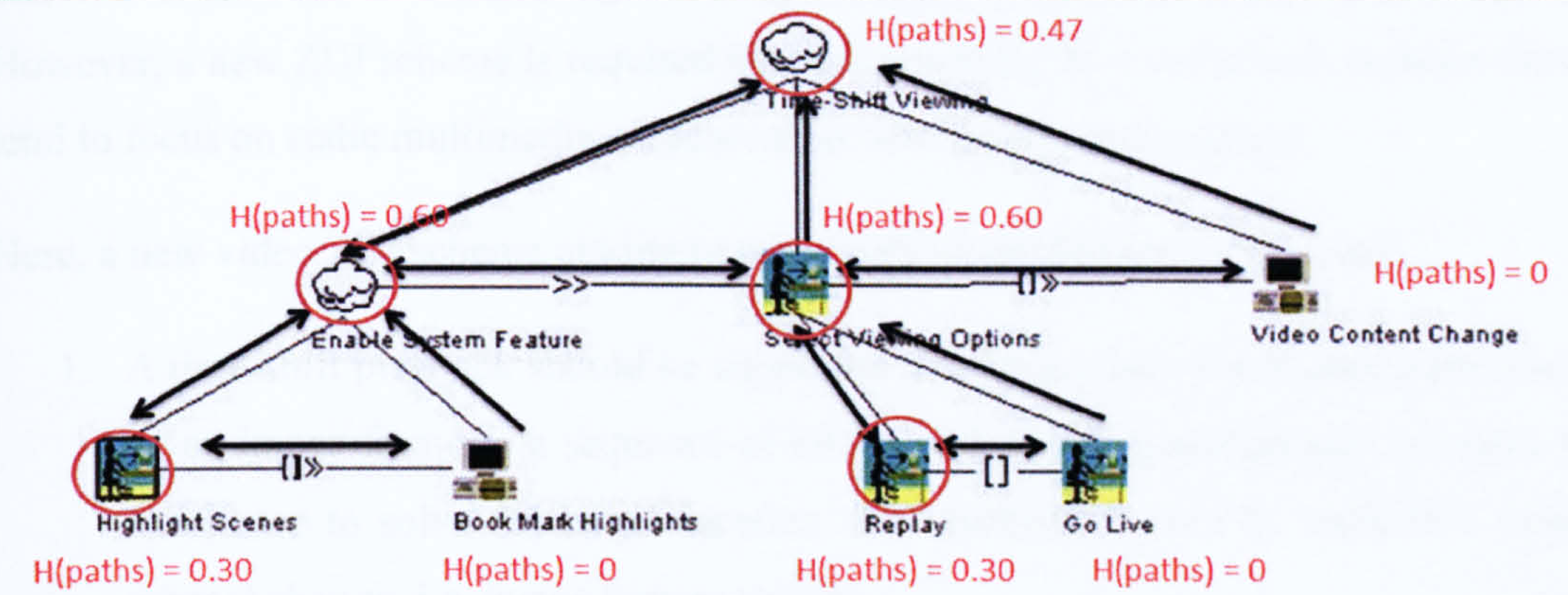
$AND\ AG(task = T2\text{-}1) \Rightarrow E((transition = ())\,U\,(task = T2\text{-}2))$

T2-2 as Initial Task:

$AG(task = T2\text{-}2) \Rightarrow E((transition = ()\,)\,U\,(task = T2\text{-}1))$

$AND\ AG(task = T2\text{-}2) \Rightarrow E((transition = RETURN)\,U\,(task = T2))$

Personalisation Implication



3.4 Task Interface Requirements for Supporting Personalisation

As discussed in section 2.3.4, passive adaptation can effectively filter video content, which can be perfectly applied to sports events selection for personalisation. When constraints such as latency and network connection speed are critical to the interactive tasks such as in tasks of multi-angle viewing, selective target zooming and time-shift viewing, a passive adaptation approach may not be viable as it requires users to choose from the recommended content before the personalisation goes into effect. With this consideration, the task user interface requirements which are prerequisite to support the personalisation of interaction are defined as below.

3.4.1 Interface Requirements in Multi-Angle Viewing Task

Existing technology can be 'fragile' when used under live sports events settings. (See section 2.3.3) In this task, two requirements are defined:

1. It should use real camera views rather than using a virtual camera approach for multiple views of the sports events. This is because virtual camera cannot create the true viewing angles and most importantly visual distortion could reduce the visual quality of the video content.
2. A multi-stream adaptation scheme is required so that is able to adapt the multiple feeds to the available bandwidth on user terminal while preserving the visual quality is required.

3.4.2 Interface Requirements in Selective Target Zooming Task

A temporal separated zoomable user interface (ZUI) technique is a further useful requirement for live video content zooming (see section 2.3.3). In the context of live video content, e.g., live sports events, a ZUI can facilitate how users view objects of interests in different situations, e.g. viewing a long shot or using a small screen device. However, a new ZUI scheme is required to extend existing ZUI techniques because these tend to focus on static multimedia content rather than on dynamic content.

Here, a new video ZUI scheme aiming to address these challenges is proposed:

1. A time-shift playback should be supported to help increase the zooming precision on an image frame or a sequence of image frames that a user expects to zoom-in and hence to solve the target location shift problem caused by rapid live video content change, i.e. image frames change.
2. Multi-bitrate videos streams should be used for live video content zooming. This is inspired by approaches used by existing zooming applications with other multimedia content, e.g. Deep Zoom [3] uses multi-resolution images to achieve a high frame-rate visual quality experience.

3.4.3 Interface Requirements in Time-Shift Viewing Task

Based upon lessons learnt from existing editing and sports video content processing approaches, as discussed in section 2.3.3, a live time-shift directing system is required to have the following two interface requirements:

1. System based editing: live editing requires the system to provide users with easy editing options as simple as allowing replay or not.
2. Use of a human director's expertise: To achieve the first requirement, the system should also be able to embrace formal director's rules complied from observing the work of human directors.

3.5 Summary

In this chapter, the interaction requirements for a next generation A-V player were presented. The main tasks supported are to select events from the current sports events, to support multi-angle viewing, time-shift viewing and selective target zooming. In order to model user interactions and to personalise an interactive system, a new interactive task

[3] http://msdn.microsoft.com/en-us/library/cc645050%28VS.95%29.aspx

model is afterwards proposed. The proposed model not only addresses the challenge of 'annotating traceable user actions' as discussed in section 2.2.2, but also enhances existing task models in other ways. In addition to that, the proposed personalised interactive task model is able to identify the critical interactive sub tasks that require personalisation. Further, task interfaces that support personalisation are defined for each interactive task. Table 3-9 summarises the critical interactive sub tasks and the interface requirements.

Table 3-9 Critical interactive task for personalisation and associated task interface requirements

Abstract Task	**Critical Interactive Sub task**	**Interface Requirement for Personalisation**
Sports Events Selection	Browsing	N/A
Multi-Angle Viewing	Activate Camera Stream	Using real physical camera Multi-stream adaptation scheme
Selective Target Zooming	Zoom in	Time-shift playback after zooming in Multi-bitrates video streams
Time-shift Viewing	Highlight and Replay	System dominant editing Use of human director's expertise

Chapter 4

4 Personalisation in a Next Generation A-V Player

Personalisation concerns tailoring interactive tasks to a user or a group of users' preferences (Figure 4-1). There are two sources of user input to do this. First, explicit input can be given by the user in terms of domain specific preferences. This is usually given at the start of the first session of use but it can be designed to be updated at any time. Second, implicit user input can be acquired by monitoring and classifying a user's interaction with the system. Both explicit user input and explicit user input is used in this input.

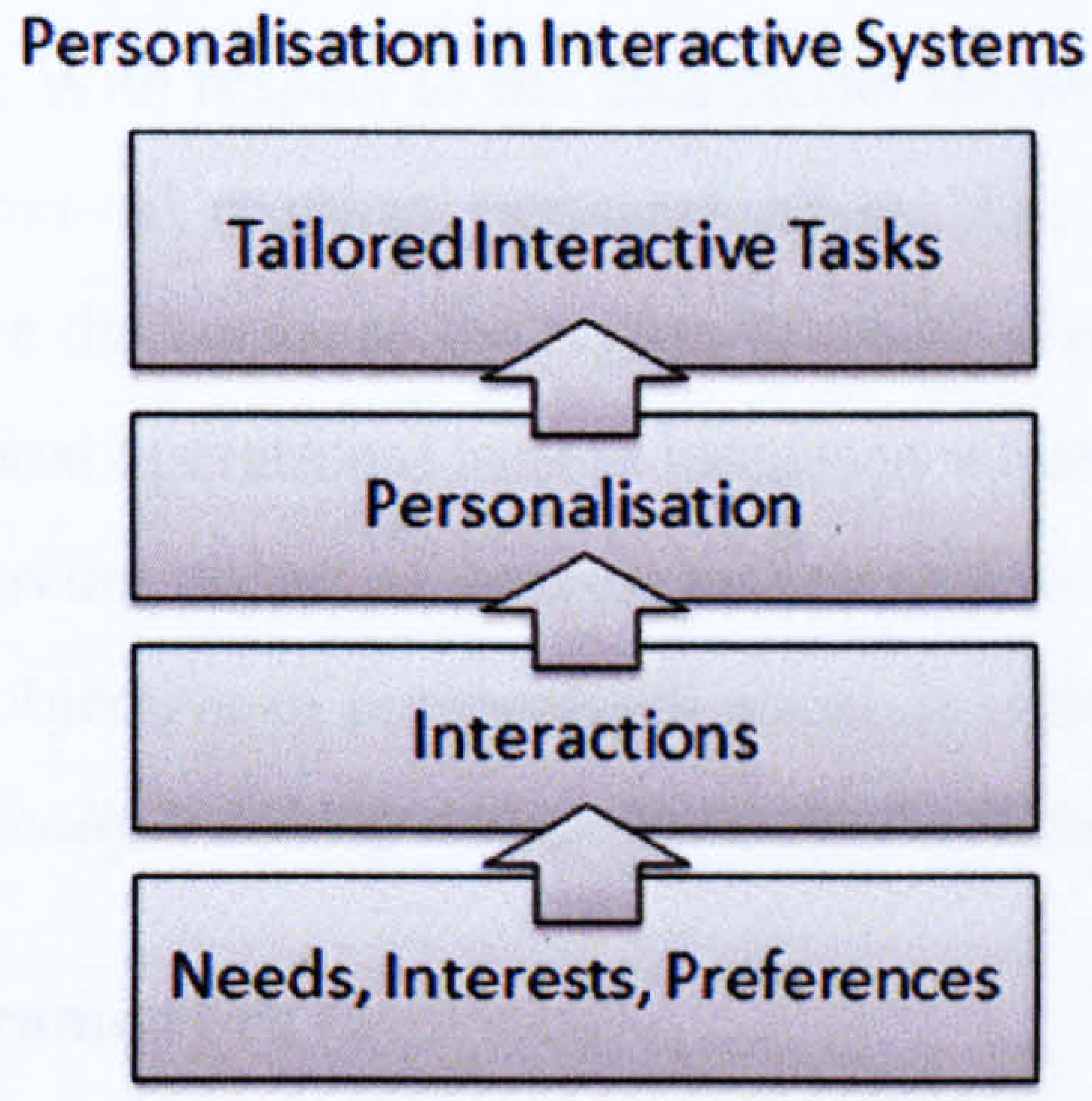


Figure 4-1 Personalisation in interactive systems

In the following sections, the personalisation objectives are presented. Then, a personalisation model that supports implicit user input is defined. Finally, this model is applied to personalise interactive tasks in a next generation A-V player.

4.1 Personalisation Objectives

The main objective of the personalisation is to reduce the load of interaction and hence to ease the interaction. In practice, each task that is personalised, based upon the interactive task models proposed in 3.2, has a task specific personalization objective.

Sports Events Selection: With respect to the task model shown in Table 3-5, the time and operational load e.g. scrolling, required for browsing will both increase when the amount of browsing content increases. In order to mitigate such operational cost, the objective of personalised browsing is defined as: *to enable users to quickly select the preferred live sports events.*

Multi-Angle Viewing: With respect to the task model shown in Table 3-6, the objective of personalised camera stream activation is defined as: *to automate the user's operations when switching or activating preferred camera streams during live sports events.*

Time-Shift Viewing: With respect to the task model shown in Table 3-8, interaction during the sub-tasks of '*Highlight Scenes*' and '*Replay*' require personalisation. These sub-tasks are traditionally performed by a human director during live events. The personalisation objective of time-shift viewing is defined as: *to enable the system to support users to highlight scenes and replay the highlights using a 'virtual director'.*

Selective target zooming: With respect to the task model shown in Table 3-7, interaction during the sub-task '*zoom-in*' requires personalisation. The zoom-in feature requires users to identify and move the cursor to the region in which objects stay in the frame. For video content, the additional operational load is increased whenever a user needs to move the cursor to follow a moving object of interest and to define the extent of the object of interest. As a result, the objective of personalised zoom-in is defined as: *to support users in selecting and automatically tracking a zooming target of interest.*

4.2 Personalisation Framework Overview

The personalisation framework (Figure 4-2) for a next generation A-V player for live sports event viewing support personalised interaction depends upon the types of interactions that are selected to be personalised and the types of context-based constraints personalisation adapts to. The framework embraces both network and terminal centric personalisation modes. Individual personal profiles can be fed into a group-based knowledge base to support scalable personalisation for live broadcast content. It is more scalable because the system in the network centric mode adapts to each group rather than to each individual. The system can further adapt if a terminal centric mode is used. One of the key design issues here is harmonising the effect of group-based personalisation versus the effect of individual personalisation.

User interaction starts with the sports event selection task which enables the other tasks including multi-angle viewing, selective target zooming and time-shift viewing. User's usage information of these tasks is maintained as part of the user profile. In addition, user's demographic information such as gender and age are also kept in the user profile.

User profiles represent a data structure that reflects users' viewing preferences during viewing sessions. Viewing parameters can be encoded and stored on a local terminal or network. The process is viewing session driven, i.e. each record item is linked to a particular session. Two crucial user operational driven parameters are dictated by the system, one is the event type and the other is the duration of the viewing session. The event type represents the user preference for the sports event type and it is obtained from the video content metadata created when the content is created. The viewing duration indicates the relative duration of a user's preference, i.e. the actual live event total duration of the event divided by the actual viewing duration by a user. This data is encoded in XML based structure such as: **<session id="zhenchenwang@eecs.qmul.ac.uk" date = "04/05/2009 21:27:36", event = "Football", duration = "0.23" />...<session .... />**. In this personalisation framework, each individual user profile is bound to a unique user ID and password registered with the system. Here, a user's email address is used as a user ID.A newly registered user will get a system generated password that is sent to this email address. After logging in to the system through a log in screen, a particular user's profiles of his/her past operations can be tracked, retrieved and updated. Individual user profiles can also form the group profile which is used in group personalisation.

Personalisation of user tasks is achieved via terminal centric processing and network centric processing (Figure 4-2).

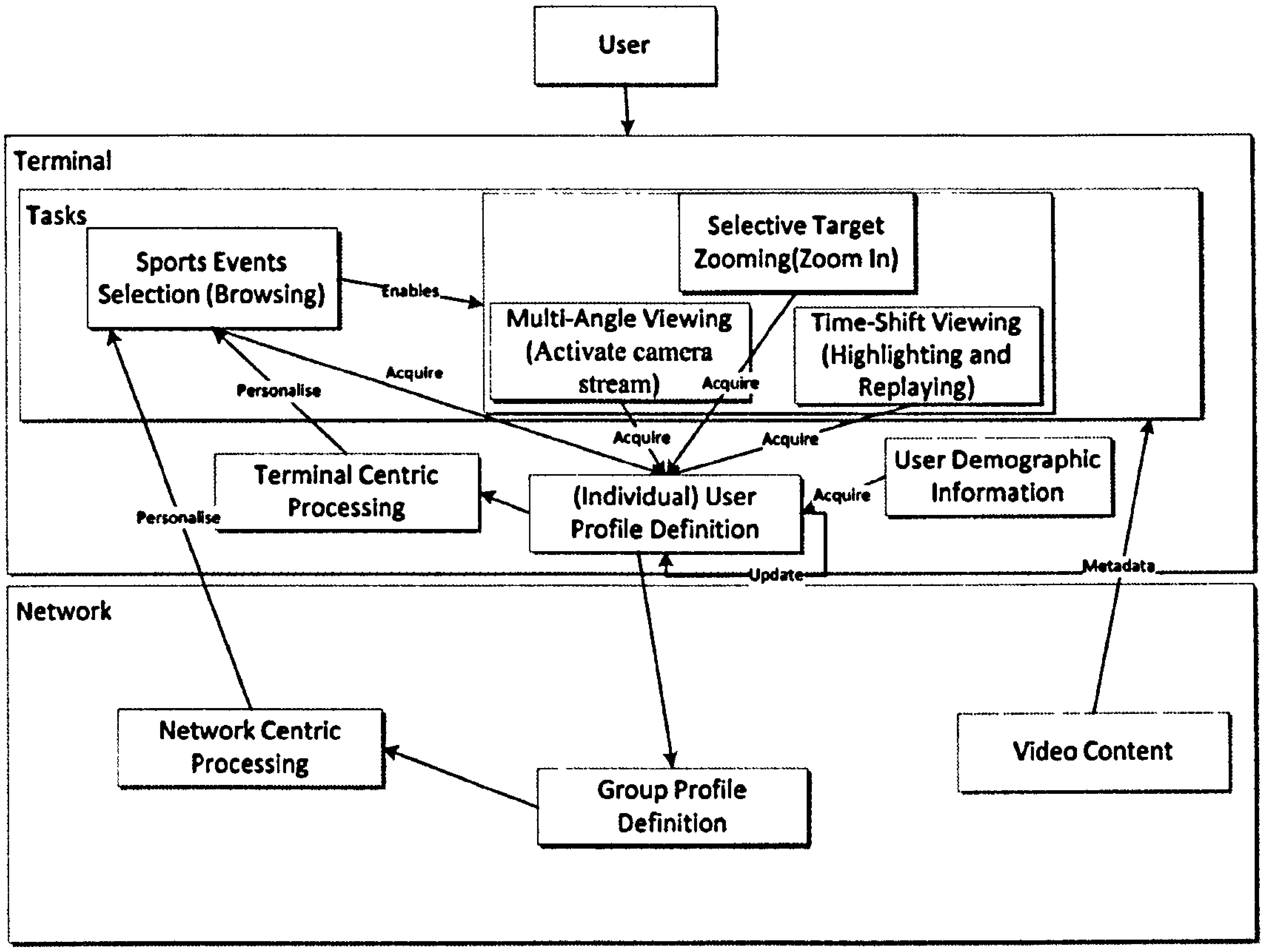


Figure 4-2 Personalisation framework

4.2.1 Personalisation Design Issues

In a personalised interactive system, some service, e.g., video content delivery, is adapted to the user model or user profile. The essence of personalisation or system adaptation to the user profile is a matching and filtering process. The service instances that match a corresponding part of the user profile can be viewed and those service instances that don't match are omitted – thus reducing the information load and the operational load.

Table 4-1 Design choices

Abstract User Task	Design Choice	Justification
Sports Events Selection	Variant user profile; Both implicit and explicit retrieval of user profiles; Passive adaptation; Group recommendations use network-centric adaptation	User preferences may change as live events progress; Basic user information such demographic information are required; User may be better able to choose amongst multiple choices; Users may have similar sports events preferences; Recommendations require group profiles.
Multi-Angle Viewing	Variant user profile, Implicit retrieval of user profile; Active adaptation; Individuals use terminal-centric adaptation	User preferences may change as live events progress; Users may not want to be interrupted when watching live events; Live content context changes can be very dynamic; Individual users may vary their viewing preferences during live broadcasts.
Time-shift Viewing	Variant user profile; Implicit retrieval of user profiles; Active adaptation; Individuals use terminal-centric adaptation	As for Multi-Angle Viewing
Selective target zooming	Variant user profile; Implicit retrieval of user profile; Active adaptation; Individuals use terminal-centric adaptation	As for Multi-Angle Viewing

Different types of context adaptation can be defined (Poslad, 2009). First, part of the user profile adaptation is invariant versus variant. Second, the adaptation depends on how the user profile or user context is generated and updated. A user profile can be generated from explicit user input, e.g., users enter their preferences or from implicit user input, e.g., usage information. Third, the adaptation depends upon if the system is active or if it is passive. Fourth, user profiles may be individual or group-based. However, dependencies exist between some of these design issues, e.g. explicit user interaction may be useful to reoccur during a session because user preferences depend upon the performance of athletes of interest during the progression of a sports event. Finally, the adaptation can be terminal-centric which requires the terminal to process the user profile or network-centric which allows more powerful backend system to process user profile.

In the following sections, the justification for the pros outweighing the cons is discussed in each case. Based upon the interactive task models proposed in 3.2 and the design factors discussed in this section, the major design choices for each task are listed in Table 4-1.

4.2.1.1 Invariant versus Variant User Profile

A user profile contains both session invariant user information and session variant information. Session invariant user information describes users in a way that is relatively static across multiple use sessions, e.g., demographic information. Session variant user information describes session specific information, e.g. session duration.

In addition, part of the profile is also variant because it is context-driven within an event or user session. E.g., a user's preferences depend upon the performance of athletes of interest during the progression of a sports event until its conclusion.

4.2.1.2 Acquiring User Profiles Explicitly versus Implicitly

Usually explicit user preferences are gathered during a single-shot interaction prior to the first session; explicit preferences may not accurately match the preferences of that a user actually views. Implicit user interaction is more likely to be gathered over multi-shot user interaction during multiple sessions and so more accurately relate to what a viewer actually chooses to view.

Explicit preferences are often retrieved in a pre-session via an explicit user interface where user preferences are input (Jameson et al., 2004). In part because sports event preferences are variant and multi-faceted, specifying the event type, event venue athlete's performance, etc., explicit preferences could also be gathered during a user session. A score voting system can be used to allow users to express their own judgements concerning athletes' performance (Schalleck et al., 2004; Van Beusekom et al., 2004) during events. Users can be explicitly asked to express preferences for pre- (prepared) event segments (Zhang et al., 2007). Explicit user profile acquisition approaches are less effective in live sports events broadcast scenarios as users' preferences may not be updated in a timely fashion unless the system continuously prompts users or allows them to spontaneously volunteer amendments. Much of this related work on explicit user preferences also tends to focus on one particular sports event, e.g. football. The main drawback of these explicit approaches if used during a session involving popup questionnaires and voting is that it could possibly interrupt and distract users' viewing experiences.

Implicit preferences can be obtained via the monitoring of user interaction with the system (De Ávila and Zorzo, 2009). Attention analysis can be used to model users, attention targets including directors, the audience and commentators, and to detect the event highlights (Ren and Jose's work, 2006). Implicit preferences can more naturally reflect varying contexts in the content. However, this introduces complexity, as updates need to be aggregated.

4.2.1.3 Active versus Passive Adaptation to User Profiles

In active adaptation, the system performs the adaptation on behalf of the user. For passive adaptation, the system provides information and recommends or proposes an action but the user actually chooses to perform the adaptation, i.e., to trigger an action, rather than the system. Active versus passive adaptation has pros and cons. Passive adaptation enables a user to have the opportunity to accept or reject a proposed system task but because this is manual there is a delay in the system action being triggered or not. Active adaptation in contrast can more dynamically adapt to a changing application context, e.g. by performing some of the task actions directly. By doing so, it can also substantially reduce the user operational load. A disadvantage of active adaptation is that user may have less control over the system if users are unaware of such automation (Höök, 1999).

4.2.1.4 Adapting to Group versus Individual Profiles

An individual profile represents the preferences of a single user and be performed client-side in the user access device or server side (in the network). The individual user profile can represent one single acquisition (single-shot) of explicit user preferences or be an aggregate of multi-session (multi-shot) user preferences. A group profile by comparison represents an aggregate of multiple users that are clustered into different cohesive groups where within each group, users share similar preferences. Different aggregation functions are used to generate a group recommendation or preference (Jameson and Smyth, 2007; Yu et al., 2006). Two of the most common types of aggregation function are based upon k-nearest neighbourhood approach (calculating Pearson Correlation that represents the preference data of top-N nearest neighbours of the particular user that are weighted by similarity) (Min et al., 2011) and collaborative filtering (predicting what users will like based on their similarity to other users) (Ping et al., 2009). Group user profiles are likely to be maintained server-side rather than clients-side. Individual profiles tend to relate to represent a content-based matching, whereas group profiles tend to be based upon a collaborative filtering approach.

There are pros and cons in adapting system interaction to group user profiles and group recommendations versus to individual ones. The use of group profiles has two main benefits. It solves the cold-start problem of users having to make explicit preferences about many things. Instead this relies on the greater number of preferences that arise from many users. Group profiles also solve the problem of having a system adapting the same service, i.e., content, in a many individual ways instead the number of different adaptation relates to the number of groups which is much less. The cons of using group profiles are that there needs to exist a sufficient density of users to form groups or clusters. Secondly, users need to be matched to the correct groups else a user will find the adaptation not to be useful.

4.2.1.5 Terminal-centric versus Network-centric Adaptation

Terminal-centric adaptation can be used to first build a knowledge base of individual user preferences and behaviours. This requires the use of heavy terminal processing, to support basic coarse-grained feature extraction from video content, thus making this method less suitable for deployment on low resource terminals such as mobile phone type terminals and less suitable for Web based terminals. In contrast to terminal-centric adaptation, network-centric adaptation allows user profiles to be shared with content and service providers and enables group-based user profiling using a network-centric personalisation model. Such a model supports more advanced and finely grained feature extraction from video content and supports scalable personalisation for live streamed video content to many thousands of users.

4.2.2 User Model or Profile

The first main challenge in personalisation is in acquiring and maintaining a hybrid user model or user profile using a mix of explicit user input (user preferences) and implicit user input (usage data) that can change during live events and across user sessions (Section 2.3.2). The second main challenge is in defining how specific user tasks are personalised by a system (section 3.2.2).

In this user model, the task specific objective is predetermined with respect to the type of task. Two different types of user preferences are differentiated namely invariant preferences and variant preferences (Figure 4-3).

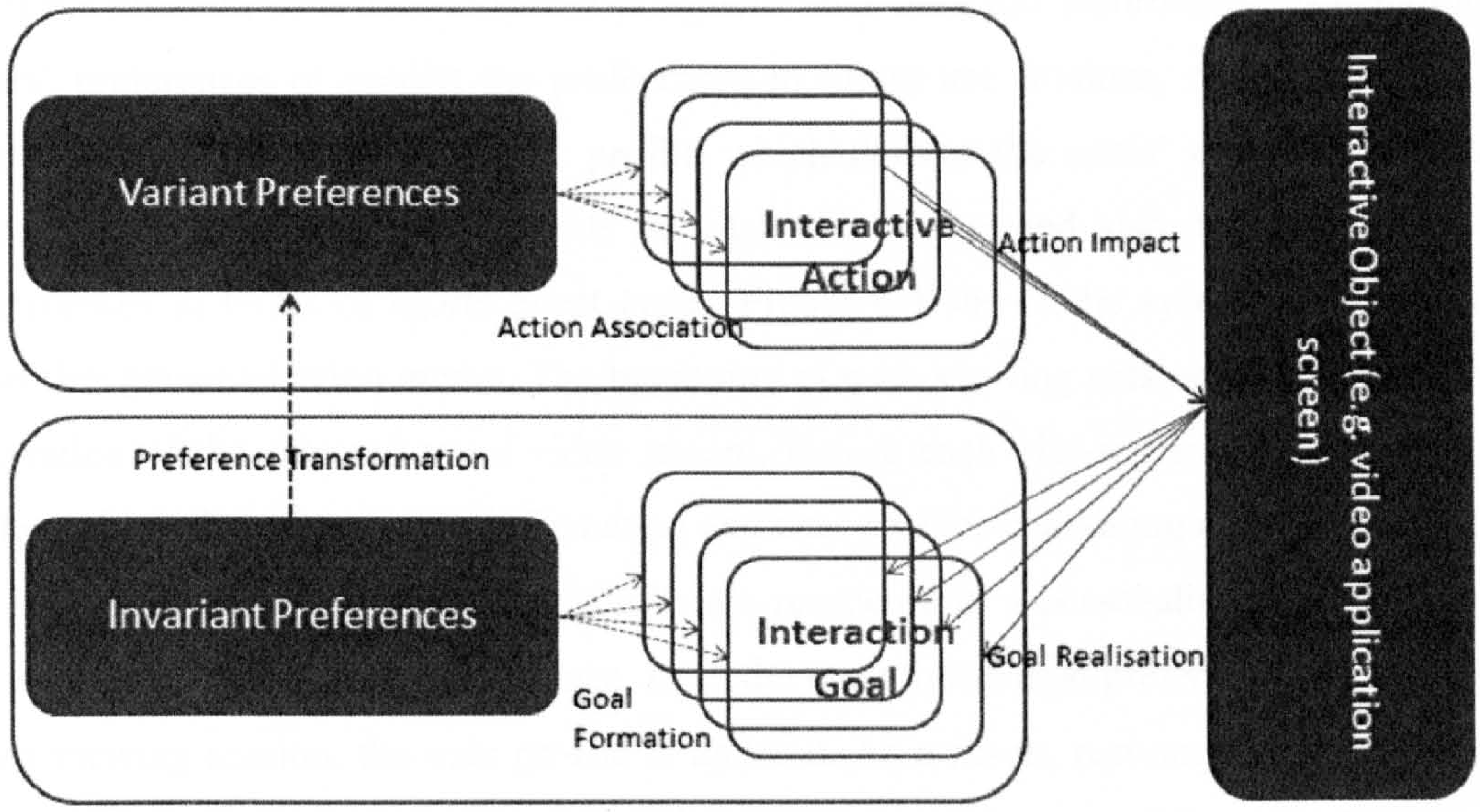


Figure 4-3 Use model in a personalised interactive system

Invariant user preferences are implied by the overall task objective, e.g., in a zoom-in task, the interaction goal is 'to zoom into the image on the video screen. An invariant preference can be to 'prefer to view the long jump take-off using a close-up'. The invariant user preferences are extended from the variant user preferences to make these achievable in the task domain by taking into account specific task constraints, e.g. 'prefer to zoom into the centre region of the video screen rather than corner of the screen because that's where the long-jump take-off board visually appeared.'

A variant preference is associated with *interaction actions* which are designed to have direct impact on the interactive object, e.g. 'move the mouse cursor, click mouse left button.' An *interaction object* receives a user's interactive actions and produces a system response, i.e. action impact, e.g. a specific change to the content displayed, and leads to the realisation of the *personalised interaction objective*, e.g., the magnification level increased with the magnification focus being on the take-off board.

In this thesis, the variant preference is particularly concerned. This is because in the live sports events context, user's preferences are more dynamic as sports events progress.

4.3 Personalised Event Selection

Personalised sports events selection ranks a list of live events in accordance with user preferences of event types. Users initially select events in order to trigger the system to start the creation of a user model. The system later uses this information to determine users' preferences or predict the preferences in future use sessions. A user's model is maintained in the form of a user profile which defines the users' observed viewing durations for particular sports events for each use session and also defines user event preferences in terms of sports event types. Figure 4-4 shows the system sports events selection personalisation model. The beginning of each viewing session is defined as the activation of the selected event video stream. Before each viewing session, the system collects both live sports event information and user profiles containing usage information of previous viewing sessions. The results are reordered in a personalised event list that ranks sports event types by preference, from the most to the least preferred. At the end of each viewing session, the user profile is updated. As a result, recommendations can be obtained via both user profile and defined preference factors of live sports events. Preference factors that affect this model include the importance of the events, the type of events and the participants of the events. Among these three factors, the event type is the key part of a user preference while the other factors are used to sort the final personalised event list in a descending/ascending order in terms of their value, e.g. events in their final round will be ranked as the first one for a particular event type when multi-round events exist.

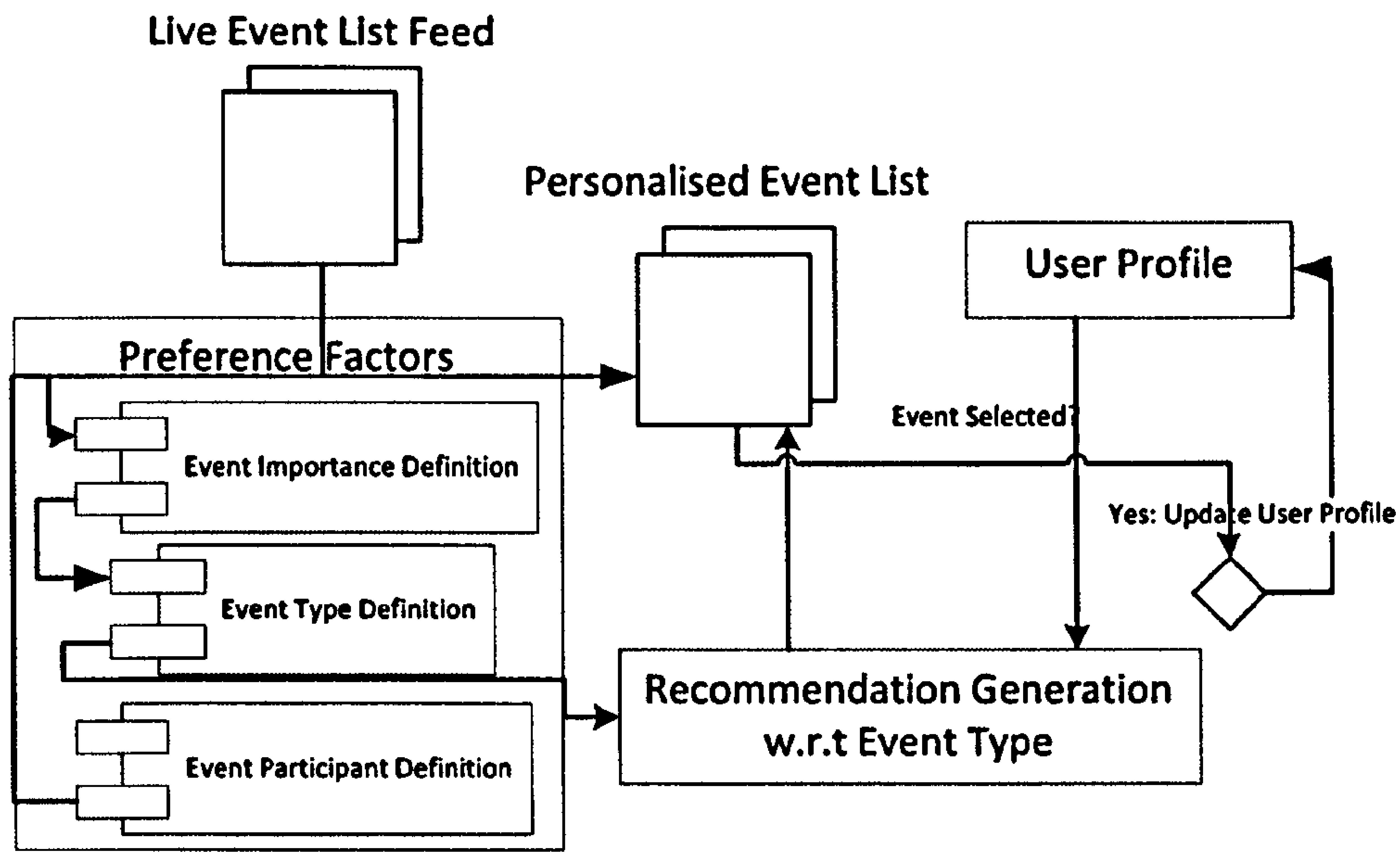


Figure 4-4 Sports events selection personalisation model

There are several key challenges to personalising event selection. During major sports events, multiple sports events instances occur concurrently. Viewers may also have multiple preferences. Viewers' preferences for viewing sports may be finely grained and multi-valued. Rather than select sports to be viewed just by sub-genre, additional multiple dimensions, can be used for selecting content, e.g., by the type of competitive round such as preliminary versus final and by demographic characteristics such as nationality. However, even matching participating athletes' nationality to a viewer's nationality is more complex than it seems. For example, someone may be of Chinese descent but may be a national or a long term resident in another country, e.g., England. It is not clear whether such a viewer would prefer viewing Chinese or UK athletes participating in specific events.

Human viewers have one focus of attention to view multiple events even if a multi-view screen is used. For live multi-sports events, an individual viewer's schedule is semi-deterministic and is dynamic. A viewer can decide to subjectively switch between events because of several reasons such as the current view is not captivating, the preferred athletes are not successful as expected, because of specific event incidents or because of the score status, etc. Note also that live multi-view sports channels are dynamic reflecting the event schedule. There may be no fixed channels as in TV entertainment systems so one can't always identify a specific sports view by number. Hence, video recommender systems tend to recommend channels by programme content genre rather than by channel number. In some research cases, the content and preferences can be represented semantically. Semantic descriptions may alleviate the issue that users must know a priori the content provider descriptions, rather they can use their own semantics for their preferences and match these to the semantics of the content provided. Note also, that for video retrieval, quite detailed explicit user feedback can be gathered from users during the interaction. However, explicit feedback from users during viewing of live events needs to be minimised else the usability of the system is reduced as it detracts viewers from the immersion of following an event.

A further complexity is that viewers may wish to mix and match live event viewing versus interacting with near live event views, e.g., slow-motion replays of the end of a competitive event. It is also challenging to generate metadata for the dynamics of live events in real time, in contrast to the off-line generation of metadata for video retrieval, and to design system support to semi-automate view switching, e.g., using recommendations. Whereas a majority of a potential live event audience often appears to want the convenience of viewing a time-shift or delayed events, because of other commitments such as daily activities, work etc., a significant minority, up to a quarter of an audience, often prioritises viewing the event as it happens – live. Some event instances have a semi-deterministic scheduled start, and duration, depending on the location, event organisation, duration of athlete warm-up, participation, post-event activities, the success and the ranking of athletes in the event instance and any abnormal event incidents.

As a result, personalisation is needed that handles:

1) Concurrent live multi-sports events that are constrained by a dynamic schedule;
2) Multi-valued, dynamic, user preferences (that may change as events progress);
   A characterisation of domain objects that match user selections that reflect multiple semantics rather than a universal semantics.

4.3.1 User Model

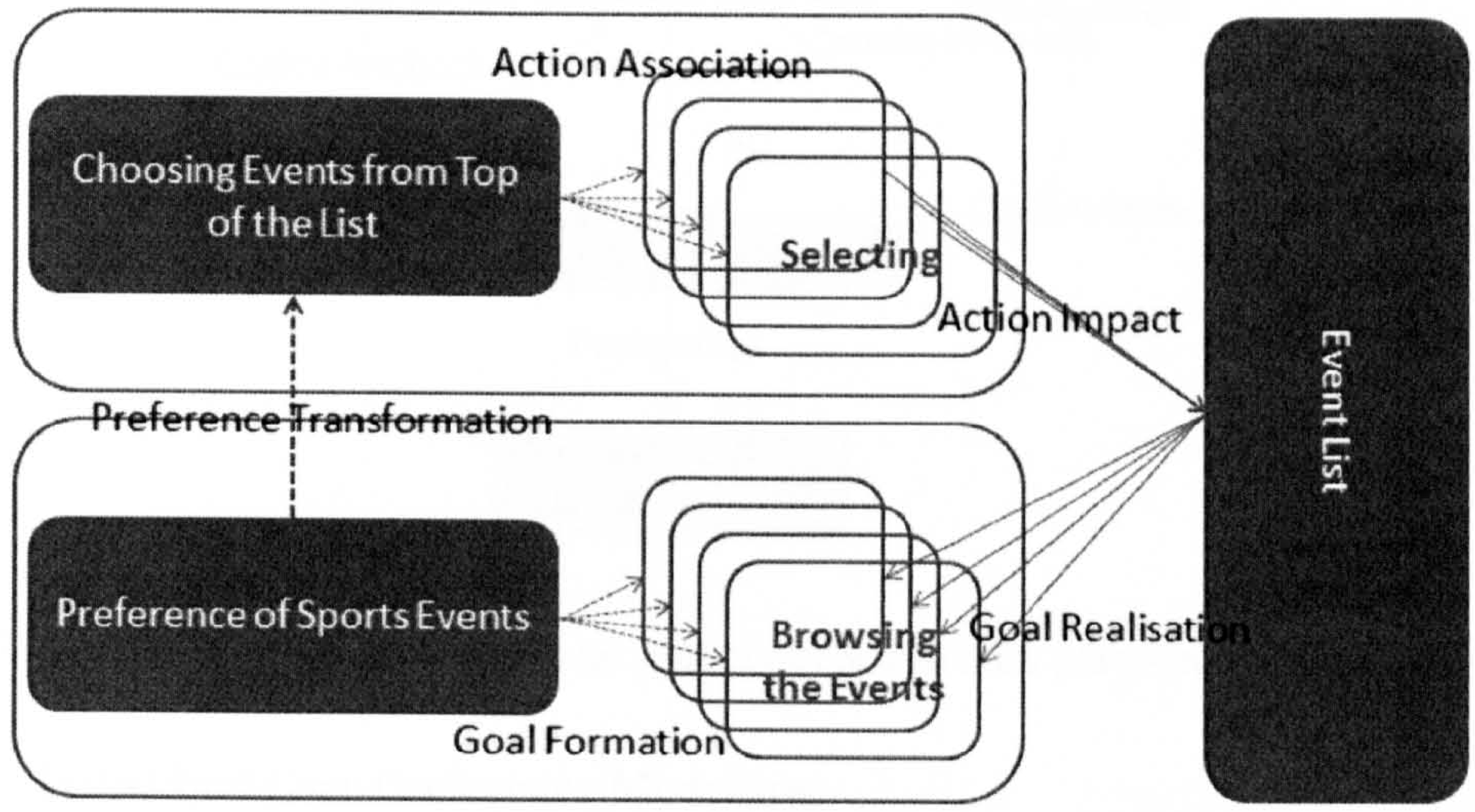


Figure 4-5 Personalised event selection user model

By reference to the user model proposed in section 4.2.2. The invariant preference of this interaction can be defined as the user preference of the sports. The variant preference will be choosing events from top of the list. The variant preference is defined in accordance to one characteristic of the list which normally requires user to scroll down to reach the preferred events. Therefore, the objective of this interaction (see section 4.1) can be achieved by minimising such scrolling efforts, i.e. operational load. In Figure 4-5, the overall user model of sports event selection is given as an application of the model proposed in Figure 4-3.

4.3.2 Sports Events Selection based upon Individual Recommendations

Figure 4-6 shows a traditional individual feedback driven personalisation model. In this model, user preference modelling is heavily dependent on explicit user feedback (Boratto et al., 2009; Masthoff and Gatt, 2006; Xu et al., 2002). When watching live sports events, this approach can be distracting for many users. In the My-e-Director 2012 field trial, less than 10% of users actually provided explicit feedback when watching sports events (My-eDirector field trial report, 2011). The personalisation model proposed in this section addresses this issue through enabling the system to implicitly acquire users' preferences.

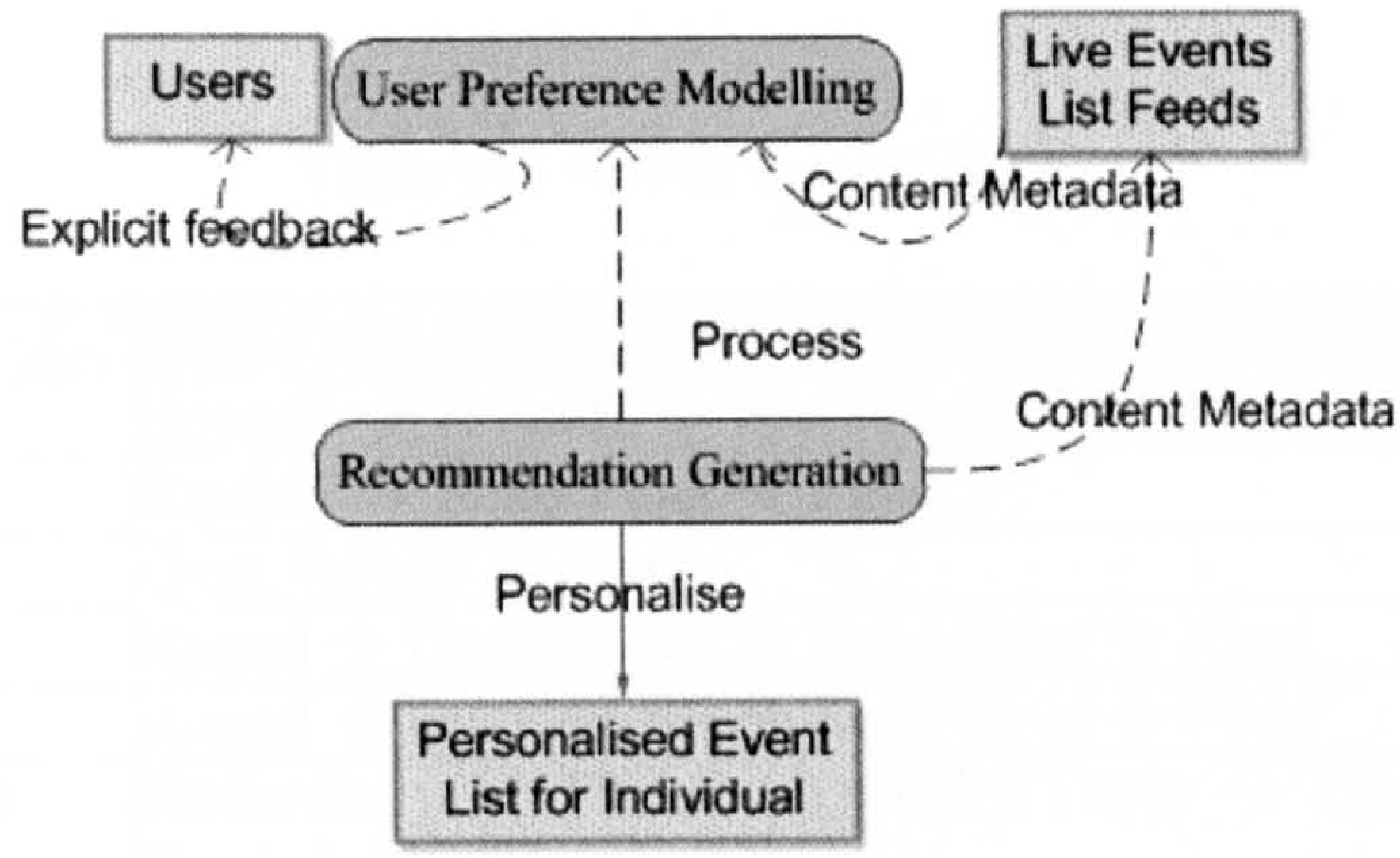


Figure 4-6 Traditional individual user feedback driven personalisation model

4.3.2.1 Individual User Preference Modelling

In this model, data that includes event importance, event participant and historical viewed sports events and the viewing length are used in the adaptation process.

Event Importance refers to the stage of the sports event. Many sports events share similar competition stages. In Table 4-2, different sports events are listed with respect to the event stages. Figure 4-7 presents an example of an events importance continuum.

Table 4-2 Sports Events Stages

Event	Stage
100m race	Round1 → Round2 → Semi Final → Final
1500m race	Round1 → Semi Final → Final
Long Jump	Qual. Round → Final
Football	Round → Quarterfinal → Semi Final → Final
Basketball	Round → Quarterfinal → Semi Final → Final
Beach Volleyball	Preliminary → Round → Lucky Loser → Quarterfinal → Semi Final → Final
Swimming	Heat → Semi Final → Final

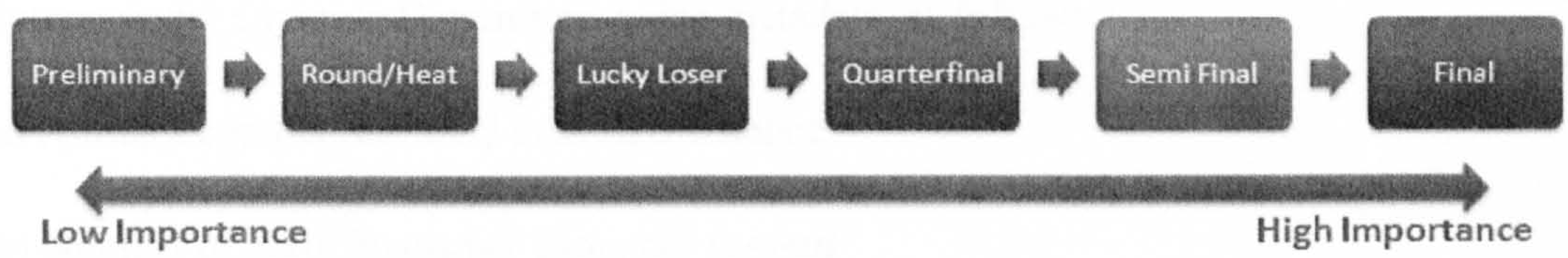


Figure 4-7 Events Importance Continuum

Event Participant indicates the performance ranking of an athlete. The participant performance index (PPI) is calculated as follows.

$$PPI = \frac{\sum_{i=0}^{n} r_i}{n}, where\ n = team/atheltes\ number, r = world\ ranking$$

The smaller the PPI the more successful participants would appear during the event. Note that the world rankings can be obtained from third parties e such as *All-Athletes.com* [4].

Event importance and event participants are defined here as a special filter for live sports events. In this thesis, it is envisioned that people tend to watch more important events and events with more successful athletes in. The first sorting process occurs after receiving the live sports events list, and the second one will be done after user preferences are matched to the initially sorted events list.

Sports attribute describes the sports discipline. There are around 26 summer and 8 winter sports disciplines according to the International Olympic Committee[5]. These

[4] http://www.all-athletics.com. (accessed Nov. 2nd 2010)

[5] Official website of the Olympic Movement, http://www.olympic.org/uk/index_uk.asp(accessed May 8, 2009)

sports can be classified according to multi-dimensional properties including their play strategies (invasive, defensive, etc.), achievable targets (gate, distance, etc.) and equipment (Gallahue et al., 2003). However, some of these properties cannot be clearly applied, e.g., football or soccer can be thought of as either an invasive or defensive game. In addition, there may be definition overlaps among these properties, e.g. football can be a ball game in terms of physical equipment but also as a field game in terms of the venue of the game. Sports events in this thesis are defined in terms of properties associated with international game rules so that each event type can be officially distinguished.

A sports event is normally shaped by a set of common game rules. These rules are expressed as different quantitative properties. Here, such properties are defined by the International Olympic Committee, using metadata, as follows:

a) Athletes competition field area size/distance

b) Number of participants per game per session

c) In-game session number

d) Minimum number of competition direction changes (the maximum number is defined to be 10 in this thesis)

e) Number of standard technical incidents

As an example, football is described with defined metadata as shown in Table 4-3. Note the order of the metadata follows the definition above. Table 4-4 lists the other events with metadata descriptions.

Table 4-3 Football Event Metadata Description

Football = {8520, 22, 2, 6,10, 22}	
a	110 x 75= 8250
b	22 players
c	2 sessions
d	10
e	Hand ball; offside; Foul; Charge; Corner Kick; Direct free kick; Dive; Dribble; Free kick; Goal; Goal kick; Indirect free kick; Kick-off; Own goal; Penalty kick; Penalty shoot-out; Penalty spot; Yellow/Red card; Save; Tackle; Take a dive; Throw-in (22 in total)

Table 4-4 Metadata Description of other Events

Event	Metadata Description
Basketball	{420, 10, 4, 3, 24}
100m race	{100,8,1, 0,5}
400m swimming	{400, 8,1, 7,2}
5000m race	{5000,16,1,0,5}
Beach Volleyball	{128,4,3, 7,10}
Long Jump	{55, 1, 3, 0, 3}

4.3.2.2 Recommendation Generation

Content-based filtering matches events from the live event lists in terms of the event type to a user profile. There are three steps to define the user preference:

- *Step 1)* At the beginning of viewing session T, calculate the user preferences with respect to the defined event metadata by using all available historical profiles before session T.
- *Step 2)* Compare the latest accumulated user preferences for live sports events from a viewing session T-1 to T-t with respect to the live sports event list for this viewing session, where t is a threshold value that is used to adapt to the user preference orientation change, e.g. events in final rounds can be more attractive, whilst neglecting past sessions.
- *Step 3)* After the viewing session T, update the stored user profile in terms of the viewed event type and relative viewing duration by appending the profile collected in session T to the existing user profile.

The sports event defined by the event metadata can be expressed in a six dimensional vector. Different events can be compared with respect to this, e.g., using two relative irrelevant events, football and swimming, a similarity value between the two events can be calculated. Given the event vectors

Ea: $(E_A = (E_AM1, E_AM2,\ E_AM3,\ E_AM4, E_AM5))$ and

Eb: $(E_B = (E_BM1, E_BM2,\ E_BM3,\ E_BM4,\ E_BM5))$, the geometric angle between two vectors can be expressed as following formula.

$$\alpha = arccos(\frac{< vE_A|vE_B >}{||vE_A|| \, ||vE_B||}) \qquad (4.3.2.1\text{-}1)$$

The smaller the angle between them, the more similar the two sports events are. As a user's preferences are defined in terms of event metadata. Hence, user A's preference can be represented as

$$U_A = (U_AM1, U_AM2, \ U_AM3, \ U_AM4, U_AM5, U_AM6)) \qquad (4.3.2.1\text{-}2)$$

U_AM1 is known as user A's preference of first metadata type, i.e. competition field area. The user preference of metadata instances can be accumulated over a predefined number of viewing sessions. As a result, the formula 4.3.2.1-2 can be transformed to:

$$U_A{:}\left(U_A \frac{\sum_i^n M1 * P_i}{n-i}, U_A \frac{\sum_i^n M2 * P_i}{n-i}, \dots, U_A \frac{\sum_i^n M6 * P_i}{n-i}\right), \qquad (4.3.2.1\text{-}3)$$

$$where \ P_i = \frac{t_i}{T}$$

i denotes the beginning of ith viewing session which starts with a new list of live events, n denotes the end of the nth viewing session. Each viewing session represents a switch of event viewing so that the difference of $n - i$ would be the number of viewing sessions that are predefined. P_i denotes the duration of the ith viewing session as a percentage of the total viewing duration of all sessions, i.e. session duration t_i divided by total viewing duration T.

The vector values for a user preference are normally different from those pertaining to the current sports event vector, even if they share a common pre-defined sports metadata description, unless a user keeps viewing one particular event. At the beginning of a new viewing session, the latest user preferences vector are compared with the current sports event one using formula 4.3.2.1-3. The comparison produces a ranked event list that is resented to the user.

4.3.3 Sports Events Selection Based upon Group Recommendations

The use of group recommendation is advantageous over individual recommendation as discussed in 4.2.1.4 particularly when there are many users or much video content exists. A direct aggregation of a group of users' personalised event lists based upon individual recommendation models in Figure 4-6 or the model proposed in last section may not be an option to turn it to a group-driven personalisation model. This is because three practical issues, relating to recommending live events to groups, can occur. These issues are: the *user group re-clustering overhead*; *inconsistent user preferences*; the *passive group classification problem.*

The *user group re-clustering overhead* refers to a group recommender system having to keep re-clustering users in order to keep the group definition 'fresh' because group preferences keep changing as live sports events progress and as the events schedule unfolds. User groups are often dynamic during live events. Users' preferences can be athlete-centric, e.g., viewers prefer to see their favourite athlete win or lose, or sports incident-centric, e.g., viewers prefer to see any athlete that succeeds during points scoring incidents rather than those that do not.

*Inconsistent user preference*s occur when user preferences do not have a consistent semantic meaning across different users or are not consistent as events progress. For example, if a rating score of 4/5 represents a defined preference then it could mean impressive to some users but not to others and such a preferences may change as events progress. This problem can be attributed to both an individual's subjective opinions which by nature are dynamic and multi-semantic characterization of the domain objects.

The *passive group classification* problem occurs when a user group classification relies only on explicit users' feedback such as ratings, comments, etc. However, it may be difficult to keep such group profiles fresh in a live sports events scenario.

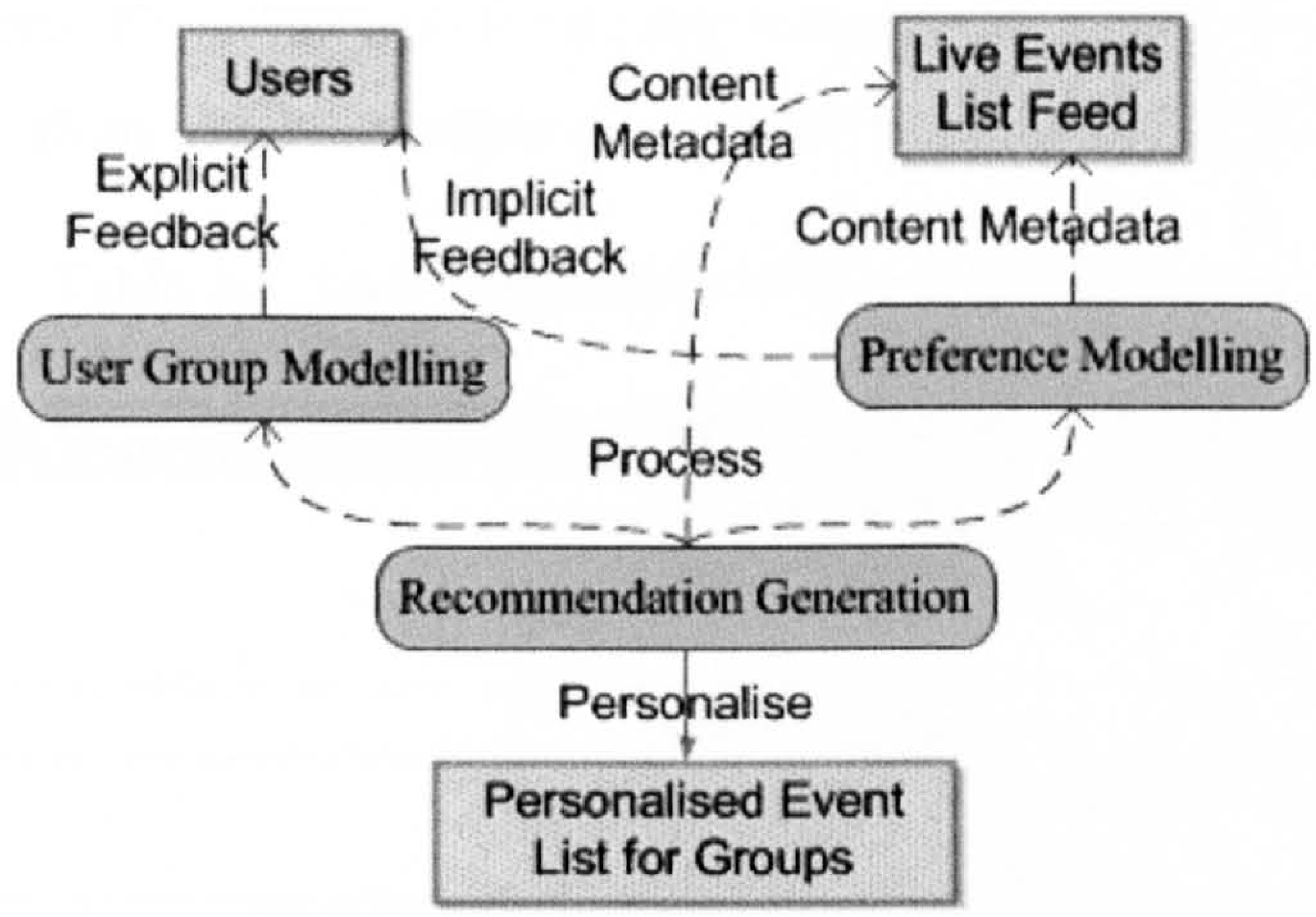


Figure 4-8 Group driven personalisation model

In Figure 4-8, a new recommendation model is proposed. The rationale of this model is as follows. First, user groups can be created with a 'one size fit all' manner. This thus reduces the group re-clustering overhead. Second, user preference models are based upon historical viewed events rather than the more subjective user rating feedback. This allows user preferences to be used more consistently across different users. Thirdly, the recommendation generation in this model is independent of user's feedback which enables active recommendations. The model also allows the system to both explicitly (in a single shot manner) and implicitly (in a multiple shot manner) acquire user profiles. The group personalisation model contains three main tasks, user preference modelling, user group modelling and recommendation generation which are further discussed in the following sections.

4.3.3.1 User Group Modelling

The advantage of using session invariant user information to group users is twofold. First, once users are grouped, groups can be reused across use sessions and across different access systems. Second, no system inference process is required as users can be explicitly asked to provide this information. As a result, the user group can be expressed as

$$Group_i = Agg\{\{ui_{1j}, \dots, ui_{1l}\} \in ui_1, \dots, \{ui_{ij}, \dots, ui_{im}\} \in ui_i\} \qquad (4.3.4\text{-}1)$$

ui_i denotes the ith user information and ui_{ij} denotes the jth information value of ui_i.

In this model the user demographic information including age, gender and race are used. The second category is the sports behaviour related information including the frequency for watching sports TV. Table 4-5 lists the information used and the corresponding predefined values from which a number of groups of users can be defined.

Table 4-5 User Information with predefined values

Common User Information	Predefined Values
Age	Young (<35), Middle Age(>35 and <60), Old (>60)
Gender	Male, Female
Race	American, African, Asian, Australian, European
Watching Sports TV Frequency	Daily, weekly, monthly, yearly

4.3.3.2 User group Preference Modelling

When the recommendation target is a group of users, data including user demographic information and the currently viewed sports events are used the adaptation process. Here, the sports events attributes are further modelled which extends the use of these attributes in making recommendations for individual scenarios.

User preferences can be described as a function of fondness of an item and preferences only exist when their counterpart items are chosen. i.e. $Pref(x) = f_{likeness}(x, L_x)$, where x denotes the item and L_x denotes the item list where x is chosen from. The fondness degree can vary between items based upon the value of a preference function.

A list of items can be described as an aggregation of a set of attributes with predefined values, e.g. a sports event with competition type=team, individual, stadium size=large, small. i.e.

$$L_x = Agg\{\{a_{1i}, .., a_{1k}\} \in a_1, ... \{a_{ni}, .., a_{nk}\} \in a_n\} \quad (4.3.4\text{-}2)$$

a_n denotes the *nth* attribute and a_{ni} denotes the *ith* value of the attribute. Therefore, the user preference of an item attribute value can be a function of a fondness degree of a predefined attribute value, i.e.

$$Pref(a_{ni}) = f_{likeness}(a_{ni} \in a_n) \quad (4.3.4\text{-}3)$$

A user's preference for an item attribute thus can be described as using a sort or ranking function:

$$Pref(a_n) = Sort(Pref(a_{n1}), .., Pref(a_{ni})) \quad (4.3.4\text{-}4)$$

A list of items can also be further defined as a sort function for item attributes:

$$Pref(L_x) = Sort(Pref(a_i),..,Pref(a_n)) \tag{4.3.4-5}$$

Based upon the sorted preference of each attribute value, the preferred items can be arranged in order from the item list L_x to enable a top-k group recommendation. e.g., if the $Pref(L_x)$ = team size with a large value, then teamwork events will be initially selected and ranked before individual events. Eventually the teamwork events in small stadia will be ranked before the teamwork events in large stadia. Sports are defined with six attributes defined section 4.3.2.1. The corresponding attributes values are redefined in Table 4-6.

Table 4-6 Sports attributes and values

Sports Attributes	Predefined Values (Compared to mean of a group of events)
Competition Area Size	Small size (<mean), Large size (≥mean)
Number of Players	Small number of players(<mean), Large number of players(≥mean)
Number of In-game Sessions	Small number of sessions(<mean), Large number of sessions(≥mean)
Competition Direction Change Times	Small number of change times(<mean), Large number of change times(≥mean)
Number of standard technical incidents	Small number of incidents(<mean), Large number of incidents(≥mean)

According to Table 4-6, the fondness of a particular attribute values a_{nk} can be expressed in a binary form, as defined in equation (4.3.4-10), where TRUE indicates a large quantity condition, e.g. large size competition area, large number of in-game sessions etc. whereas FALSE indicates a small quantity condition. Therefore, if let a_{nk} denotes a 'large number of players', then a Boolean value of TRUE can be assigned to the a_{nk} of current viewing event when it satisfies the TRUE condition in equation (4.3.4-10).

$$f_{likeness}(a_{nk} \in a_n) \tag{4.3.4-10}$$

$$= a_{nk} > \frac{1}{m}\sum_{j=1}^{m} a_{nkj} : True ? False,$$

$$where\ m\ denotes\ m_{th}\ live\ sports\ events$$

4.3.3.3 Recommendation Generation

The recommendation generation task associates a user group with a preference model. It also generates recommended items to new users.

The association between groups and items can be either strong or weak. Therefore, according to equations (4.3.4-1) and (4.3.4-2), the group-item association degree function is defined as sort function demographic user information to item attribute value association degrees.

$$f_{association}(Group_i\,, L_x\,) = Sort(f_{association}(Group_i \in \{ui_1, \dots, ui_i\}, \{a_1, \dots, a_n\})\,) \tag{4.3.4-6}$$

As a strong association between demographic user information and a particular item attribute value indicate the preference for that attribute, hence

$$\mathrm{Pref}(\mathrm{Group_i}, \mathrm{a_{ni}}) = \mathrm{f_{association}}(\mathrm{Group_i}, \mathrm{a_{ni}}) \tag{4.3.4-7}$$

Because a recommended item is comprised of a set of attributes, the preference of an ideal item can be expressed as the ranked preference of its attributes as

$$Pref(Group_i, Item_{ideal}) = Sort(Pref(Group_i, a_{n1}), \dots, Pref(Group_i, a_{ki})) \tag{4.3.4-8}$$

In order to recommend the top-k items to a user usr_i belonging to a group, equation (4.3.4-6) can be re-defined as a sort function

$$Pref(usr_i \in \,Group_i\,, L_x\,) = \,Pref(usr_i, L_x|Item_{ideal}) \tag{4.3.4-9}$$

$Pref(usr_i, L_x|Item_{ideal})$ denotes the user group preference of a list of recommended items given an ideal item, e.g. a 5 people with 3 in game sessions, 2 attack direction change game.

The association function proposed can be explicitly defined in terms of existing machine learning techniques. In this thesis, three well known techniques are employed, the Decision Tree (DT), Bayesian network (BN) and Bayesian Point Machine (BPM) are used to validate the model respectively.

4.3.3.3.1 Recommendation Generation using Decision Trees

This is composed of three elements: a decision node, an edge (links between nodes or between nodes and leaves) and a leaf. Here, the decision node is the common user information attribute value, the leaf is the fondness of a sports attribute value in a form of binary values, e.g. '1' means like whereas '0'means not like.

In order to choose the best attribute as the root of a decision tree or sub decision tree, information gain based upon the Shannon entropy is used to discriminate each decision node. The information gain between an invariant user information attribute value and the fondness of a sports attribute value can be expressed as

$$Info\ Gain(u_i, a_{nk}) = Info(u_i) - Info_{a_{nk}}(u_i) \qquad (4.3.4\text{-}11)$$

$$Info(u_i) = -\sum_{i=1}^{m} \frac{fq(Group_n, u_i)}{|u_i|} log_2 \frac{fq(Group_n, u_i)}{|u_i|} \qquad (4.3.4\text{-}12)$$

$$Info_{a_{nk}}(u_i) = \sum_{a_{nki} \in [0,1]} \frac{\left|u_{i\,a_{nki}}^{\,a_{nk}}\right|}{|u_i|} Info\left(u_{i\,a_{nki}}^{\,a_{nk}}\right) \qquad (4.3.4\text{-}13)$$

$Group_n$ denotes a user group set, $fq(Group_n, u_i)$ denotes the number of u_i type users in the user group class. $u_{i\,a_{nki}}^{\,a_{nk}}$ denotes the user information attribute value for which a value of either of 0 (FALSE) or 1 (TRUE) for a sports attribute a_{nk}as expressed in a fondness function . When the information gain ratio is required in the DT algorithm C4.5, so called split information can be used.

$$Split\ Info(u_{i,}a_{nk}) = -\sum_{i=1, a_{nki} \in [0,1]}^{m} \frac{\left|u_{i\,a_{nki}}^{\,a_{nk}}\right|}{|u_i|} log_2 \frac{\left|u_{i\,a_{nki}}^{\,a_{nk}}\right|}{|u_i|} \qquad (4.3.4\text{-}14)$$

$$Gain\ Ratio(u_{i,}a_{nk}) = \frac{Info\ Gain(u_i, a_{nk})}{Split\ Info(u_{i,}a_{nk})} \qquad (4.3.4\text{-}15)$$

The tree built allows each of the user information attribute values to be associated with the fondness of a corresponding sports attribute values. In order to recommend the top-k items to a new user usr_i belonging to a group, three additional processes are required, namely classification, preference ranking, and recommendation.

A classification process enables the system to find out the decisions D on fondness (i.e. preference) of each sports attribute value corresponding to a set of user information attribute values. A most-fit strategy can be used for the case that not all user information attributes values can be classified using the tree built. A preference $Pref(usr_{i,}a_{nk})$ can be obtained through examining reduced user information given a user information reduction function f_r, e.g. ignore particular demographic user information attribute value, say, and gender. The reduction function is iterated until a decision is reached.

$$Pref(usr_{i,}a_{nk}) = f_r(D) \tag{4.3.4-16}$$

A preference ranking process allows the system to assign a weight ω to the fondness of each sports attribute value. The ω will enable a ranking process for each attribute value for an attribute. The attribute value with the largest weight will be chosen as the decisive attribute value. A further ranking process will rank the attributes according to their decisive attributes' weights. Given ω as the decision accuracy, the sort function in equation (4.3.4-8) can be transformed to support descending sorting as follows:

$$\omega(usr_{i,}a_{nk}) = \frac{Pref_{True}(usr_{i,}a_{nk})}{Pref_{Total}(usr_{i,}a_{nk})} \tag{4.3.4-17}$$

$Decision_{Total}$ denotes the total number of existing data with the required user information attribute values in the classification process. $Decision_{True}$ denotes the number of correct decisions in the training data.

Finally, the recommendation process generates the top-k list of recommended sports based upon the result of equation (4.3.4-8) that gives a user group's preference for an ideal recommended sport. The preference of a list of recommended sports can be obtained through iterative comparisons between recommended sports based upon the ideal sport. This can be expressed as:

$$Pref(usr_i, L_x | Sp_{ideal}) = \tag{4.3.4-18}$$
$$Iterative\ Sort(a_{nk}^i \in Sp_i, \ldots, a_{nk}^t \in Sp_t)$$

a_{nk}^i denotes the *i*th sport's decisive attribute value a_{nk}, $Sp_{ideal} = \{a_{nk}, \ldots, a_{ot}\}$ denotes the ideal recommendation with a particular order of attribute values with the first attribute value has the highest preference priority.

4.3.3.3.2 Recommendation Generation using a Bayesian network

A Bayesian network is a probabilistic graphical model which can be used to handle uncertainty. A directed acyclic graph (DAG) links nodes with edges. The numerical component represents the conditional probability distribution of nodes in terms of parent nodes. Naïve Bayes is a simple form of Bayesian network which has one root node (the unobserved node) and assumes child nodes (the observed nodes) are independent of one another.

Here, the graph model can be structured as shown in Figure 4-9. The conditional probability between a parent node and a child node can be calculated given a training data set. This can be expressed as:

$$P(N_{Child} \mid T) = \frac{P(T \mid N_{Child})\, P(N_{Child})}{P(T)} \qquad (4.3.4\text{-}19)$$

N_{Child} denotes a predefined value of either a user profile attribute or a sports attribute. T denotes the total evidence for a child node in the training set.

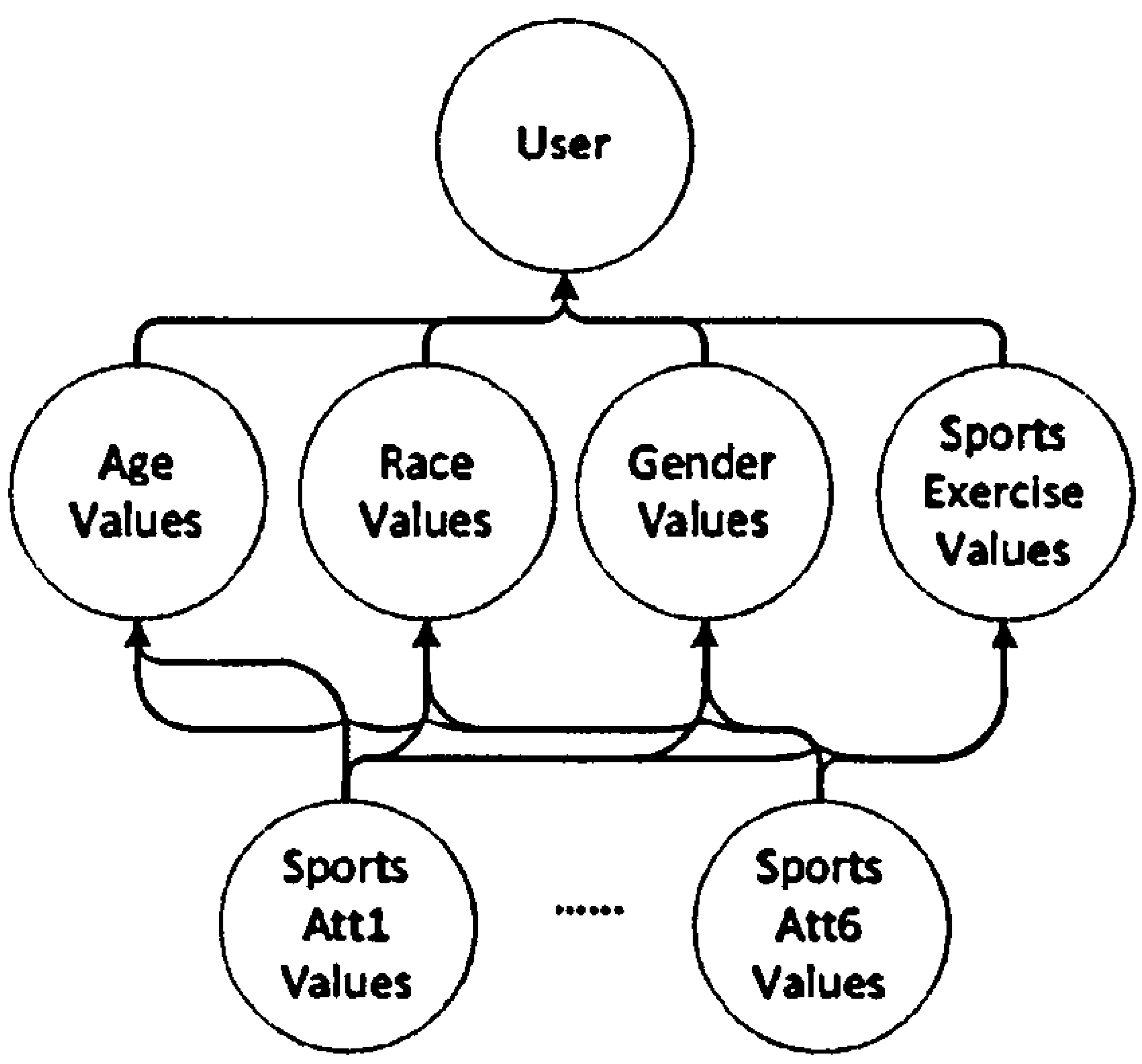


Figure 4-9 Bayesian network

Once the Bayesian network is quantified, it will be able to recommend the top-k sports to a new user. The recommendation process is similar to that of the DT approach. The only difference lies in the first step, in which a combined demographic user information (see Table 4-5) probability for each sports attribute value is obtained. The probability values are used to represent the $Pref(usr_{i,}a_{nk})$ in terms of a preference probability for each sports attribute value as follows.

$$Pref(usr_{i,}a_{nk}) := P(a_{nk} \mid usr_i \in Group_i) = \frac{P(u_1 \mid a_{nk}) \ldots P(u_k \mid a_{nk})}{P(usr_i)} \qquad (4.3.4\text{-}20)$$

4.3.3.3.3 Recommendation Generation with Bayesian Point Machine

A Bayes point machine is a learning algorithm for kernel classifiers which approximates the Bayes-optimal decision by the centre of mass of a version space (Herbrich et al., 2001). Given the hypothesis space H and the user groups training set G, the version space can be defined as

$$V(G) = \{h \in H \mid h(Group_i) = a_{nk}\} \tag{4.3.4-21}$$

In order to classify a group of users to each sports attribute value, i.e. either 0 or 1, the Bayes classification strategy is used to obtain a loss incurred by each hypothesis h applied to $Group_i$ and to weight it according to its posterior probability $P_{H|G}(h)$. The tested sports attribute value with the minimum expected loss will be chosen as the user group preferred sports attribute value. The Bayesian point algorithm thus can be defined as:

$$A_{bp}(G) = MinP_G\left(P_{H|G}\left(los\left(h(G), H(G)\right)\right)\right) \tag{4.3.4-22}$$

$A_{bp}(G)$ denotes the Bayes point which is the classifier $h_{bp} := A_{bp}(G) \in H$.

Once the group preferences are classified in terms of sports attributes values, the top-k sports recommendation can be performed following the same steps as those used in the DT approach except that in BMP approach $Pref(usr_{i,}a_{nk})$ will be linked to the posterior probability $P_{H|G}(h)$, i.e.

$$Pref(usr_{i,}a_{nk}) = P_{H|G}(h) \tag{4.3.4-23}$$

4.4 Personalised Multi-Angle Viewing

4.4.1 Multi-Angle Viewing User Model

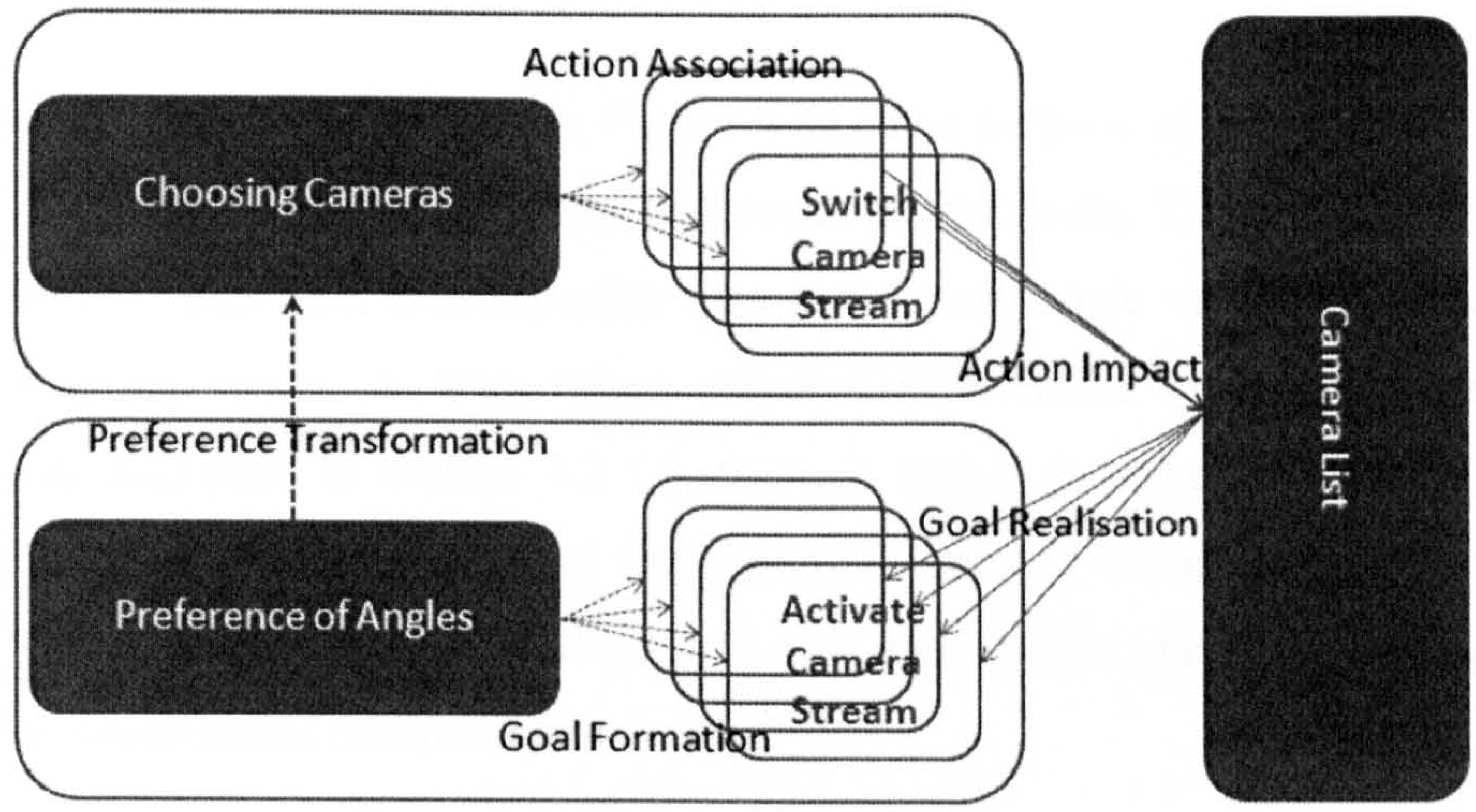


Figure 4-10 Personalised multi-view angle viewing user model

By reference to the user model proposed in section 4.2.2. The user model for multi-view angle viewing is shown in Figure 4-10.

The main objective of this interaction (see section 4.1) is to semi-automate the camera switching efforts to match the camera-switching preferences of a viewer. Viewers' preferences for camera views depend upon the type of sports event, e.g. swimming can use a camera below athletes, the patterns of incidents in that event that may use different viewpoints, the camera types and the types of viewpoints that different cameras can support.

A camera switching sequence represents how a user views a flow of incidents within an event. The switching interval represents the user's preference for a particular viewpoint.

The cameras used in sports events are defined based upon position, viewing angle and focus point. For example a fixed camera shows the athlete's foot on the take-off board in a long jump event; a high camera (e.g. on the roof of the stadium) provides an alternative view of the starting line in track events etc. The viewpoint represents the output of the camera. Different cameras may have different viewpoints or the same viewpoint.

Explicit acquisition of viewer preferences for camera views requires asking users for these. An implicit approach involves monitoring the user interaction to switch cameras.

4.4.2 Multi-Angle Viewing Personalisation Model

Personalised multi-angle viewing is modelled in terms of the interaction and activation of the camera stream.

4.4.2.1 Overview

The system architecture is shown in Figure 4-11. The purpose of this architecture is to illustrate the relationships among the five major components. These are a terminal, a Web-based user interface, a user profile data store, multi-angle viewing personalisation and a streaming server. A Web based player user interface uses a network centric approach as discussed in section 4.2.1.5, through which the other components can be linked. The Web interface receives the requested camera streams from a streaming server. This steamed content is adapted to the user profile data by the multi-angle viewing personalisation component.

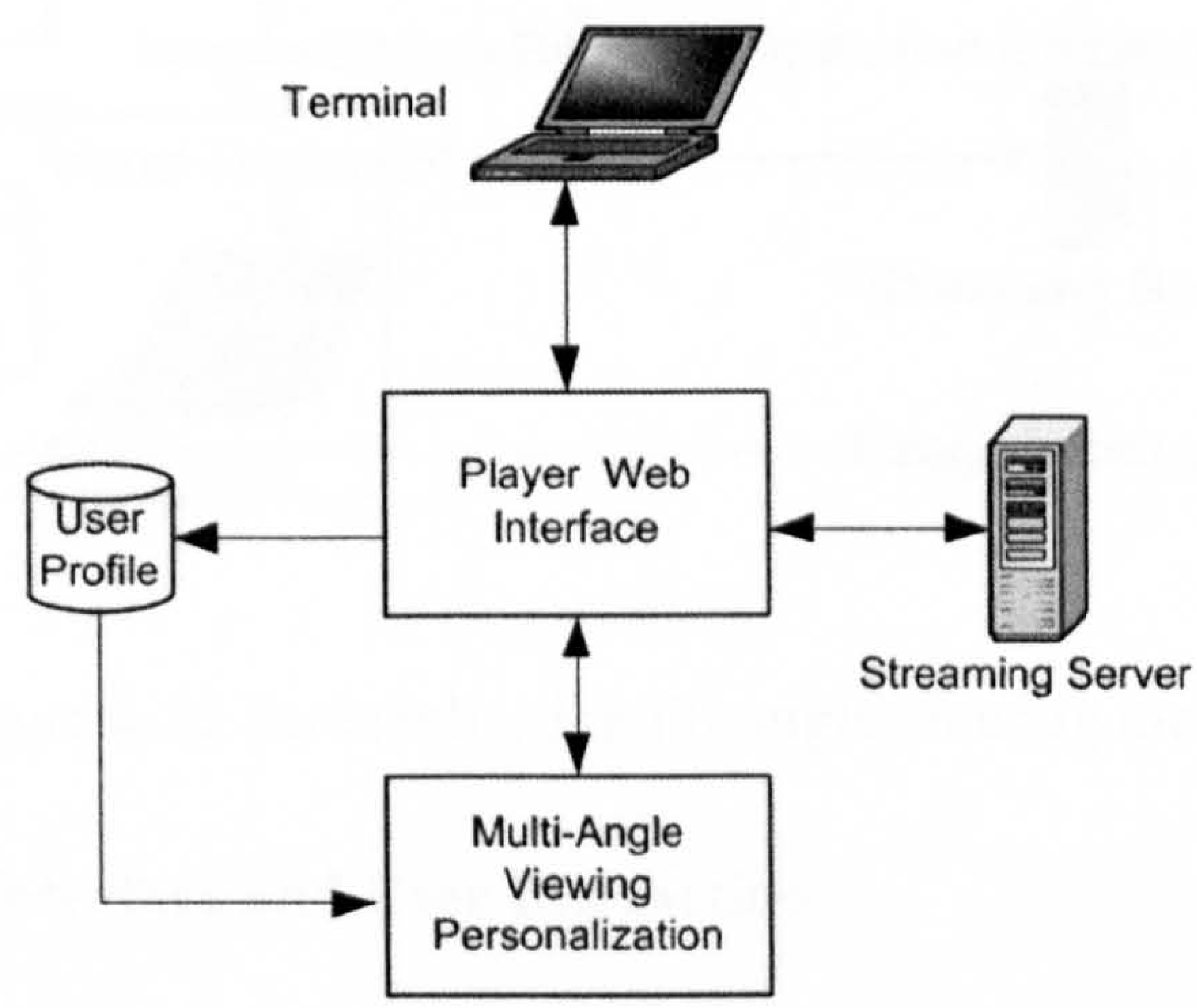


Figure 4-11 System architecture for a personalised multi-angle viewing system

A more detailed design of the personalisation model is shown in Figure 4-12. The model describes the relation between the personalisation process, user interactions, multi-stream adaptation and user profiling. Users can enable or disable the personalisation. In either case the multi-stream bitrate adaptation the user profile keeps being updated. Such a design allows the system to implicitly and continuously update a user's profile so that more historical user data can be used in the personalisation process. Each of these components is described in more detail in the next few sections.

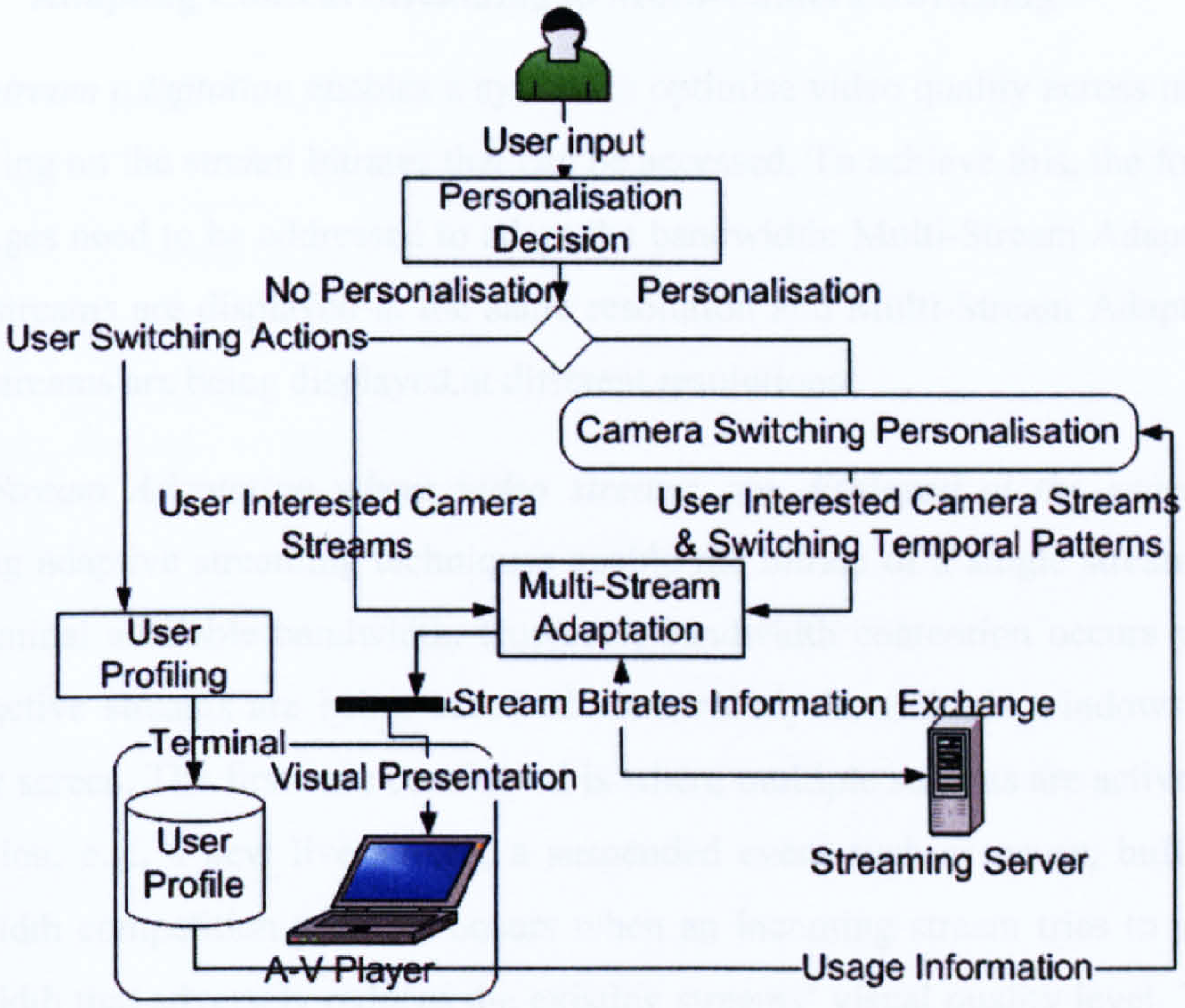


Figure 4-12 Personalised multi-angle viewing model

4.4.2.2 Task User Interface and User Interaction

User input, either a personalisation request input or a non-personalisation request initiates a multi-angle viewing Web interface. A non-personalisation request allows a user to perform two tasks: select the list of cameras of interest and switch cameras. A personalisation request only concerns the camera selection task. The function of personalisation decision making process component is used to distinguish these two types of user input by monitoring the user's choice. (e.g. manual switching or auto switching)

A personalisation request will trigger the camera switching personalisation control for user preference adaptation. The personalised switching sequential patterns and user selected camera views are used as two variables in the multi-stream adaptation process. For a non-personalisation request, there will be no user preference adaptation but a multi-stream adaptation process. For both requests, the user input information is recorded in the user profile.

4.4.2.3 Adapting Content Streaming to Multi-Camera Switching

Multi-stream adaptation enables a system to optimise video quality across multi-streams depending on the stream bitrates that can be accessed. To achieve this, the following two challenges need to be addressed to adapt the bandwidth: Multi-Stream Adaptation where video streams are displayed at the same resolution and Multi-Stream Adaptation where video streams are being displayed at different resolutions.

Multi-Stream Adaptation where video streams are displayed at the same resolution: Existing adaptive streaming techniques enable the bitrate of a single stream to adapt to the terminal available bandwidth. However, bandwidth contention occurs when two or more active streams are being accessed concurrently in multiple windows on a single display screen. The first case considered is where multiple streams are active at the same resolution, e.g., a new live stream, a suspended event such as pause, buffering etc. A bandwidth competition problem occurs when an incoming stream tries to acquire more bandwidth that adversely reduces the existing streams' visual quality level. The result of this bandwidth competition is that the visual quality of all the video streams can become unstable. A solution proposed to address this problem is that the multi-stream adaptation takes account of the video screen resolution currently used, given the available bandwidth, i.e. the higher the video window resolution, the higher the stream bitrates. This advantage of this approach is twofold. First, the bandwidth competition problem will be lessened. Second, the bandwidth can be more efficiently allocated to streams. A conventional bandwidth adaptation approach tends to allocate the same bitrates to video streams irrespective of the size and resolution of the window in which it is being viewed. Table 4-7 shows the proposed algorithm used to adapt the stream bitrates to the video screen resolution.

Table 4-7 Screen resolution bitrate adaptation algorithm

```
Input: Video stream bitrates Bs={ B1, B2,..., Bn} where Bn> Bn-1,
        Terminal screen height Ts in pixel,
        Video screen height Vs in pixel,
Output: Adaptive Bitrate Badp
BEGIN
Recursion begin
for each item Bsi in Bs
If Bi is supported by terminal
Bsupport ← add Bi to{ B1, B2, ..., Bi}
Recursion end

BiP←Supported Bitrate per Pixel = Max(Bsupport) / Ts
Init MaxB= 0← Maximal Bitrate for current video screen;

Recursion begin
for each item Bsupporti in Bsupport
MaxB← Min(Abs(BiP*Vs- Bsupporti))
Recursion end

 RETURN Badp ← MaxB
END
```

Multi-Stream Adaptation where video streams are displayed at different resolutions: In the earlier case, views are all displayed at the same resolution. An extension of this problem is when multiple videos are displayed at different resolutions, e.g., the display consists of a central large video and one or more small side videos, or it consists of a picture-in-picture (PiP) scenario. In the simple case above, the resolution adaptation approach can equally divide the bandwidth among streams but when there are screen resolution differences between multiple streams, this approach is less effective.

Here, camera streams are classified into two types, one type is the master camera stream which is envisioned to be the view with the largest screen resolution and the other type is any supporting camera stream – any other stream with a smaller screen resolutions. The multi-stream bitrate adaptation algorithm is defined in Table 4-8 below.

Table 4-8 Multi-stream adaptation algorithm

```
Input:
    Supported video stream bitrates for master stream
      B support ={ B1, B2,..., Bn} where        Bn> Bn-1,
    Supported video stream bitrates for supporting stream
      B ssupport ={ Bs1, Bs2,..., Bsn} where     Bsn> Bsn-1,

    Current master stream bitrates Badp,
    Current master stream threshold bitrates Bthre,
    Video screen height Vs in pixel,
    Supporting video stream screen height Vsi in pixel,

Output: Updated adaptive bitrate Badp for master camrea stream
      Adaptive bitrate Badpi for supporting camrea stream
      Updated adatpive bitrate for current supporting stream  B supportstream

BEGIN
BiP<-Supported Bitrate per Pixel for supporting stream = Max(Bssupport) / Ts
Init MaxBss= 0<- Maximal Bitrate for current supporting stream screen;

Recursion begin
  for each item B ssupport i in B ssupport
    MaxBss <- Min(Abs(BiP* Vsi - B ssupport i))
Recursion end

Init Bleft  <- Bn - Badp - SumExisting(Bsupportstream ) Remaining bandwidth
//CASE ONE
IF(Bleft >  MaxBss)
   RETURN Badp <-  Badp
          Badpi<- MaxBss
          B supportstream  <- B supportstream

//CASE TWO
ELSE  IF (Bleft <  MaxBss) AND (Bn - Sum(Badpi) - Bthre  >MaxBss)
    Init lowerIndex =0
    Recursion begin
    lowerIndex = IndexOf(Badp)  in Bsupport --, lowerIndex > IndexOf(Bthre)  in Bsupport
    IF(Badp -Bsupport(lowerIndex)+ Bleft >  MaxBss)
    Recursion end
     RETURN Badp <-  Bsupport(lowerIndex)
            Badpi<- MaxBss
            B supportstream  <- B supportstream

    GOTO END
   END IF
//CASE THREE
ELSE  IF (Bleft <  MaxBss) AND (Bn - Sum(Badpi) - Bthre  <MaxBss)
   MaxBss =0 <- reset the maxium supported bitrates for incoming supporting stream
    Init lowerIndex = 0
```

```
  Recursion begin
   lowerIndex = IndexOf(MaxB_ss) in B_ssupport-1, lowerIndex>0
    MaxB_ss ← B_ssupport(lowerIndex)
   IF(B_left < MaxB_ss) AND (B_n – Sum(B_adpi) - B_thre > MaxB_ss)
  Recursion end
     RETURN B_adp ← B_thre
            B_adpi ← MaxB_ss
            B_supportstream ← B_supportstream
     GOTO END
//CASE FOUR
 ELSE
     Init B_st ← set of supported bitrates for existing supporting streams
          IF(CurrentNumberOf(B_supportstream)>1)
            Recursion begin
             For each item B_supportstream in B_supportstream
             IF(Sum(Diff{B_st})AND iteritaion++ >Min(B_ssupport)
               RETURN
                    B_adpi ← Min(B_ssupport)
                    B_supportstream ← B_st(currentIndex - iteritaion) , iteration< No.of B_st
                    B_adp ← B_thre
              GOTO END
             Recursion end
                ELSE
               RETURN B_adpi = 0 ← no supporting stream allowed
                        B_adp ← B_adp
                        B_supportstream ← B_supportstream
                GOTO END
             END IF
ELSE
  RETURN B_adpi = 0 ← no supporting stream allowed
           B_adp ← B_adp
           B_supportstream ← B_supportstream
  GOTO END
    END IF
 END IF
  END IF
END
```

4.4.2.4 User Profile Modelling

User profile processing is a process of accessing, parsing and extracting usage information. In the proposed personalised zooming control model, the user profile contains the following usage information including: the camera switching sequence, switching intervals and the sports event watched. These data are encoded in XML and used in the active adaptation process. In Table 4-9 an example is given. The user profile contains one switching session for a 100m event, the 'switches' attribute contains the sequence of a viewed camera separated by hyphens. Switching intervals (in millisecond) are separated by a colon.

Table 4-9 Camera Switching User Profile Encoded in XML

```
<?xml version="1.0" encoding="utf-8"?>
<!--Created 13:12:53-->
<Profile>
  <CS id="UserID">
    <session id="28" event="M100F" switches="6643:Top-2120:Back-1672:Side-1052:Front-2094:Top-1174:Back" />
  </CS>
</Profile>
```

4.4.2.5 Personalised Camera Switching

The camera switching personalisation model is shown in more detail in Figure 4-13. The personalisation process relies on a user profile containing historical camera switching usage information and a user's personalisation requests for camera streams of interest. With camera streams setting of interest, cameras can be orchestrated in advance. With the user profile data as a training data set, the switching intervals between cameras can be obtained via hidden Markov model.

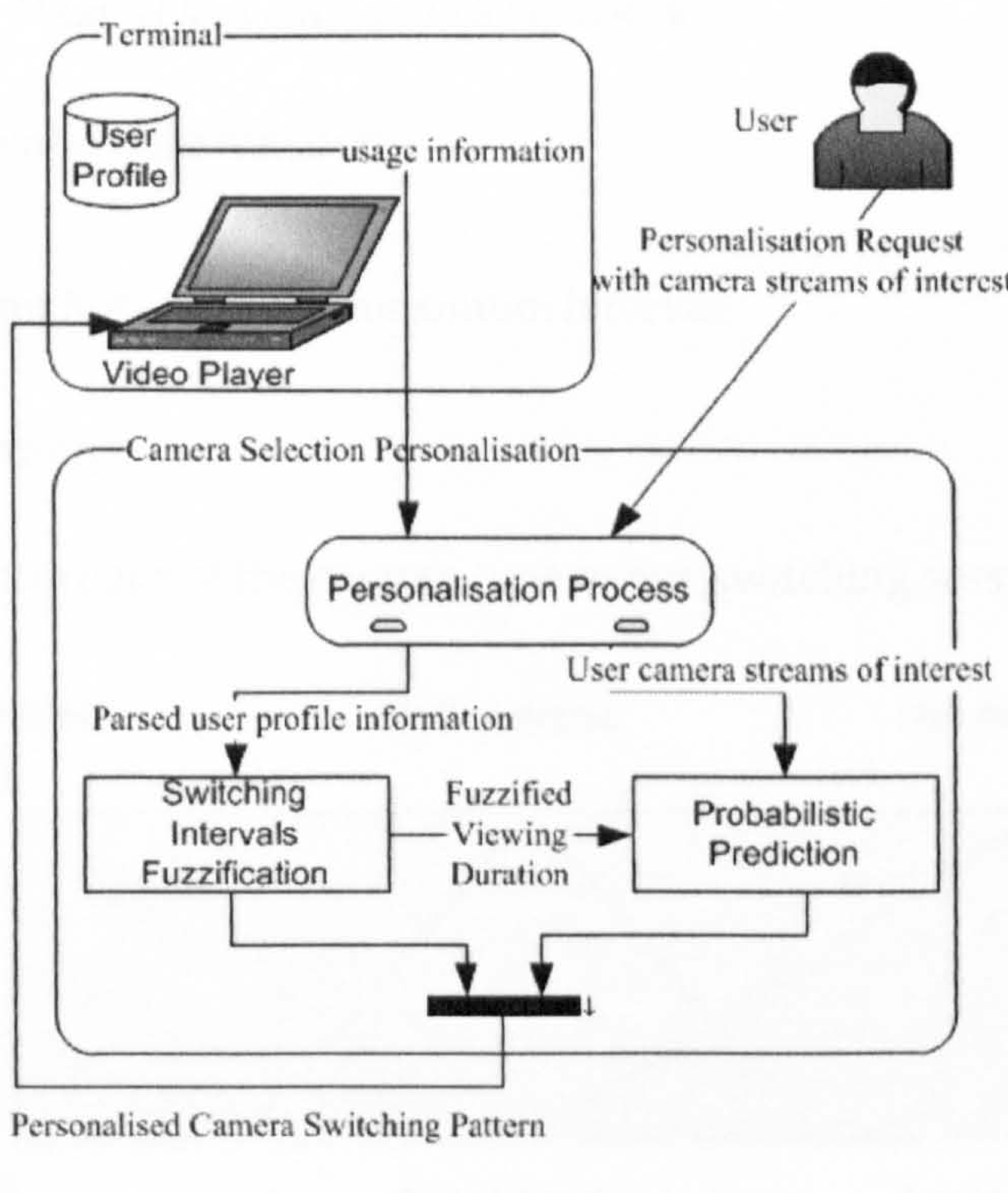


Figure 4-13 Camera switching personalisation model

The *personalisation process* subsumes two processes of a switching intervals fuzzification process and a probabilistic prediction process.

Switching intervals fuzzification: Switching intervals often have crisp values (see Table 4-9). These values change as sports events progress and when sessions are switched. Therefore it may be difficult to describe past camera switching intervals using a series of crisp values. Rather than using precise values, a more intuitive approach is to use qualitative values to describe the switching intervals such as long and short. Fuzzy logic seems applicable to describe the precise values in a qualitative form.

Here, the switching interval fuzzification process parses the encoded usage information and extracts the values of each switch corresponding to a camera type. The retrieved user profile may cover part of or all of the historical switching sessions associated with a sports event type. The interval values will be fuzzified into three membership categories, long, medium and short. The triangular membership functions ($0 < f(x) < 1$) are used to determine the degree of each membership for each crisp interval value. Three fuzzy membership functions including low degree, medium degree and high degree are used here as shown in Figure 4-14. As a result, the maximum membership value of a switching interval for a camera type determines its membership, i.e.

$$Cam_{\text{membership}} = \text{Max}(\text{f}_1(\text{X}), \text{f}_2(\text{X}), \dots, \text{f}_\text{i}(\text{X})), \text{where}$$

$$\text{i} = \text{number of membership functions}$$

$$\text{X} = \frac{\sum_{\text{i}=0}^{\text{k}} \text{D}_{\text{cam}}}{\text{K}} \text{ and X} < \textit{existing} \text{ maximum interval,}$$

$$\text{D} = \text{ interval duration}$$

$$\text{K} = \text{number of occurence of the camrea type in one switching session}$$

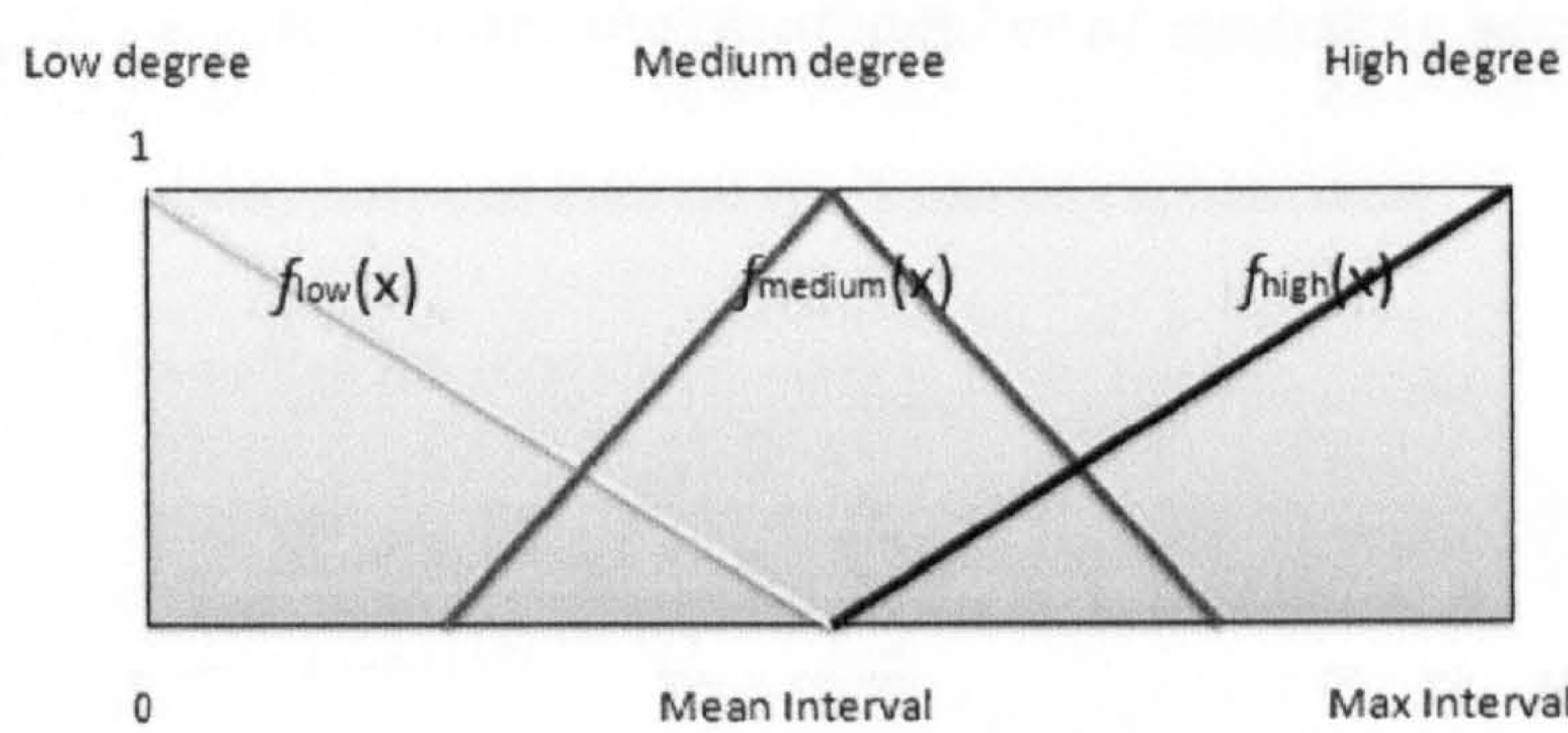


Figure 4-14 Switching intervals fuzzy sets

Probabilistic prediction: In most of the situations, it is difficult to determine how long the user would like to view an event using a particular camera view. In order to predict user's preference of switching intervals between cameras, a probabilistic model can be an option. Bayesian networks, which use graphical models to specify the conditional probabilistic dependencies between different variables, are increasingly used for the purpose of discovering some hidden user factors from a sequence of resulting data such as in Patterson et al.'s work (2003). A Markov model is investigated in many applications, e.g. a Markov model can be used to infer user intentions from sensor data (Burghardt and Kirste, 2007).

Here, a hidden Markov model (HMM) is used to predict the possible switching intervals between cameras. Obtaining the most likely sequence of switching intervals (state sequence) ***s (s1,...,st)*** corresponding to the camera sequence (observation sequence) ***q (q1,...,tr)*** with a posterior probability is a core feature of HMM. To construct the HMM, the fuzzified switching intervals and camera types are used as the state sequence and observation sequence respectively. In Figure 4-15, the topology of the HMM is illustrated and each edge represents a probability from one state to another. The proposed HMM thus can be expressed as ***M (I, S, O)*** where ***I*** = (I*i*: *i* =1,...,N) denotes the initial-state probability vector. ***S***= (S*ij* : *i,j*=1,...N) denotes the one-step state-transition matrix and ***O*** = (O*i* : *i*=1,...,N) represents the vector of a state-dependent observation distribution.

The initial switch could result in three interval types. The initial-state probability for each switching interval type can be obtained from a user's profile. This can be expressed as:

$$I_i = P(s_i = i) = \frac{\Sigma(start \to s_i)}{M}, \ \textstyle\sum_{i=1}^{N} I_i = 1$$

where D_i is ith the interval type,

M *denotes the total number of switching sessions*

The transition probability between interval types can be expressed as:

$$S_{ij} = P(s_t = j \mid s_{t-1} = i)$$

$$= \frac{\sum_{k=1}^{M}(s_i \to s_j)}{\sum_{k=1}^{M}\sum_{p=1}^{N}(s_i \to s_l)}, \quad \sum_{j=1}^{N} S_{ij} = 1, where\ l < N$$

Similarly, the transition probability between camera type and interval is

$$O_t(q_t) = P(q_t \mid s_t = i) = \frac{\sum_{k=1}^{M}(s_i \to q_k)}{\sum_{k=1}^{M}\sum_{p=1}^{N}(s_i \to q_p)}$$

$$\text{where } k, p < N$$

In order to tell the likely sequence of the interval type based upon the observed camera type switching sequence, the Viterbi algorithm can be used to maximize the joint probability ***P(s, q | M)*** in terms of ***s***. Thus the joint probability ϕ can be denoted as:

$$\phi_t\,(i) = maxP(s_{t-1}, s_t \; = i; q_t)at\; t-1 = P(s_t(i), q_t),$$

$$where\; i = 1, \dots, N\; and\; t = 1, \dots, T$$

$$\phi_{t+1}\,(i) = maxP(s_t, s_{t+1} \; = \; i; q_{t+1})at\; t = max(s_{ji}\phi_t\,(j))O_i(q_{t+1}),$$

$$where\; i, j = 1, \dots, N\; and\; t = 1, \dots, T$$

The joint probability is recursively performed with respect to *t* and will stop when

$$J = argmax\; s_{ji}\phi_t\,(j)$$

This leads to the optimal switching interval sequence up to ***t*** in state ***i,*** and a joint probability with camera switching sequence ***qt*** . With this information, state sequence ***s**** in t+1 can be obtained. $i.e. S_{t+1}(i) = (s_t(J), i)$ as well as its joint probability p^*. Table 4-10 summarizes the steps of performing the Viterbi algorithm.

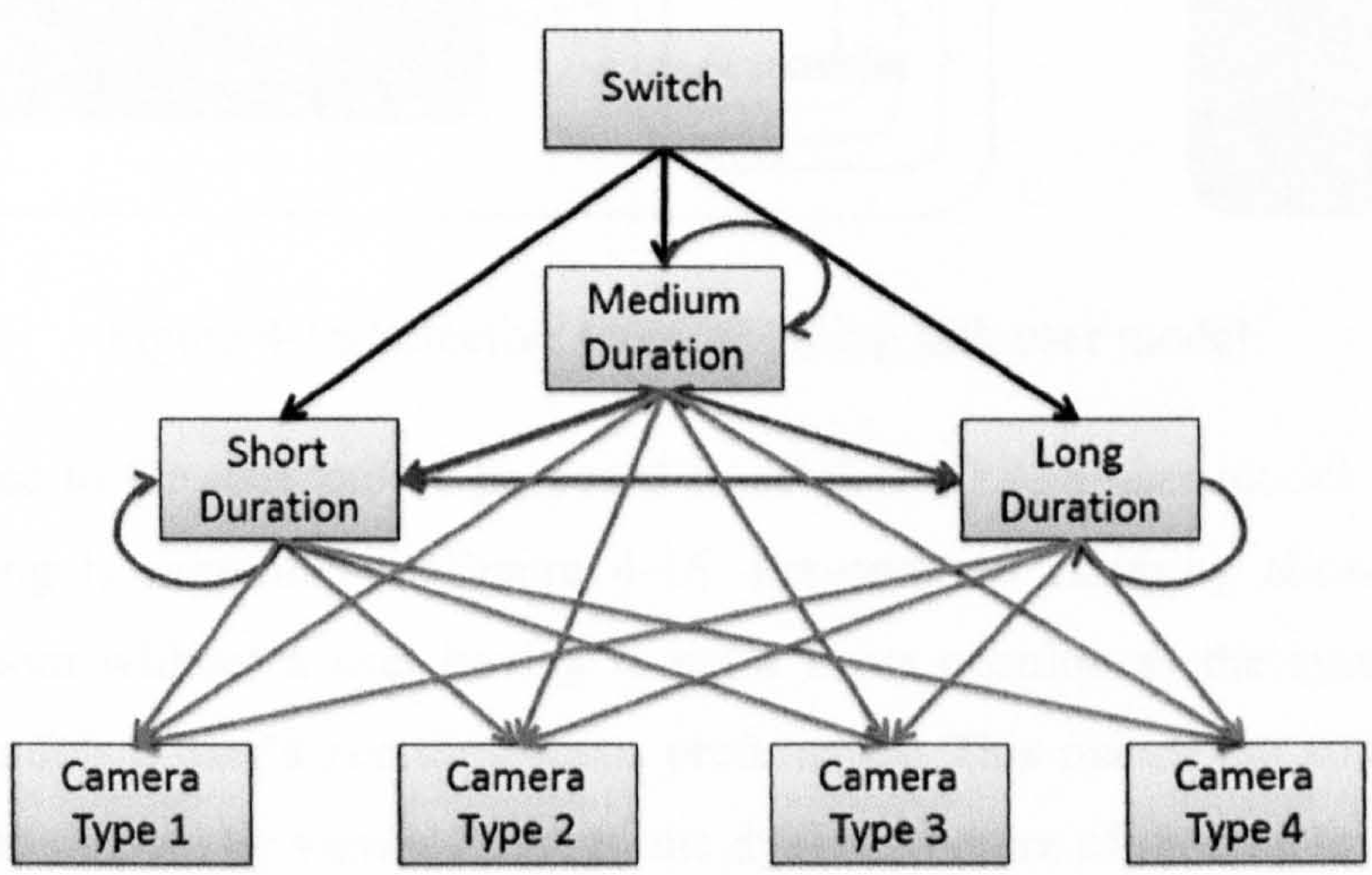


Figure 4-15 Camera switching HMM topology camera types

Table 4-10 Viterbi algorithm

Step 1:
Initialization

$$\phi_1(i) = I_i\, O_i(q_1), where\ i = 1, \dots, N$$

Step 2:
Recursion

$$\phi_{t+1}(i) = max(s_{ji}\phi_t(j))O_i(q_{t+1})at\ t, where\ t = 1, \dots T-1, i, j = 1, \dots, N$$

End when $J = argmax\ s_{ji}\phi_t(j)$

Step 3:
Result

Maximum posterior probability and state sequence at T+1

$p^* = max\ \ \phi_T(j), s^* = s_T(J)$ where $J = argmax\ \ \phi_T(j), j = 1, .., N$

4.5 Personalised Selective Target Zooming

4.5.1 Selective Target Zooming User Model

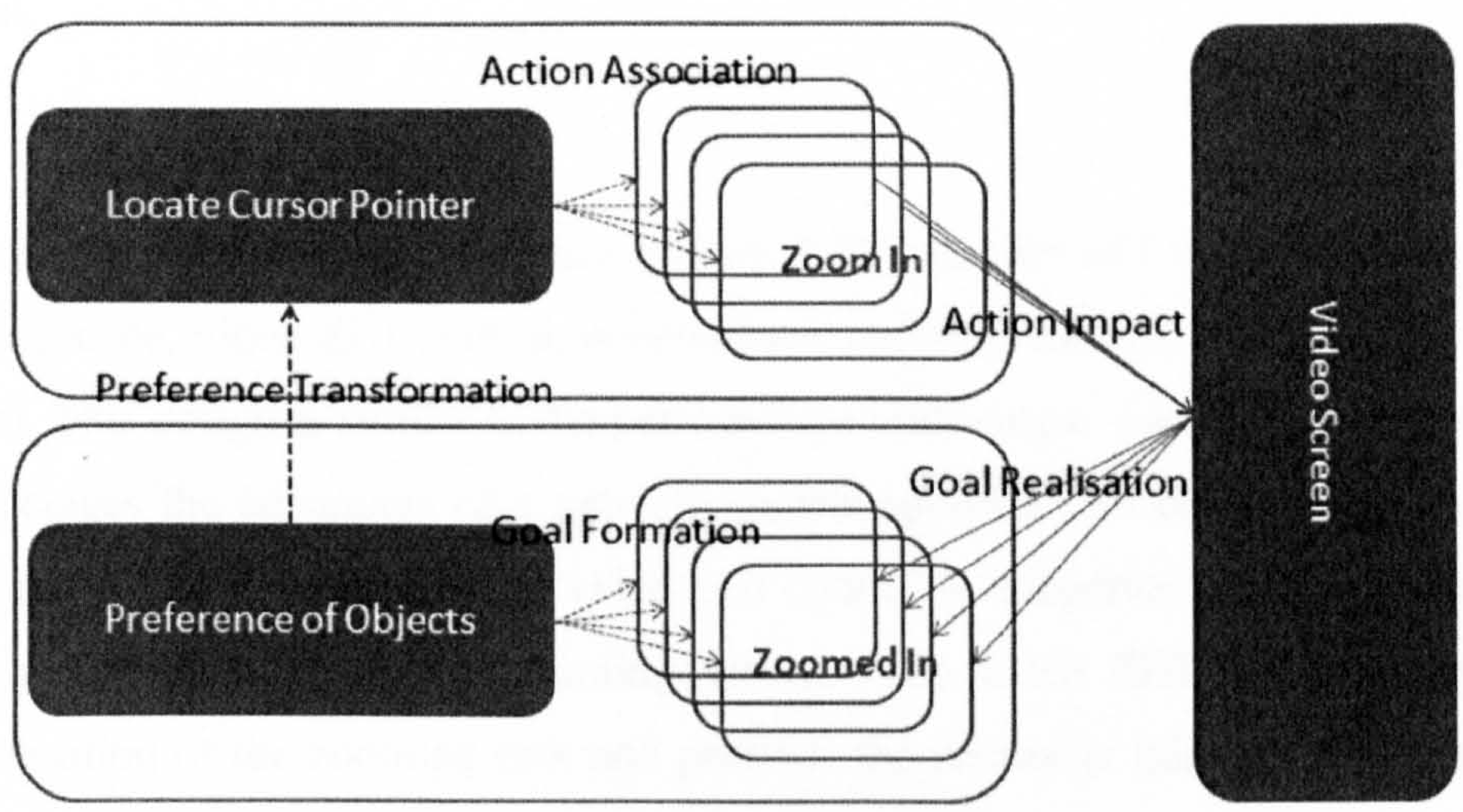


Figure 4-16 Selective target zooming task user model

With reference to the user model proposed in section 4.2.2, a user model for selective target zooming is presented in Figure 4-16. Personalised zooming allows a user to perform a zoom without a user having to set a focus position as the system sets this because it models a user's zooming focus preferences. This makes the zooming action much easier to operate by viewers. Due to the dynamic nature of moving images and the occurrence of non-deterministic incidents during sports events, user's zooming preferences need to be observed and learnt.

In order to acquire both invariant and variant preferences of observed objects in video content, 34 recorded camera feeds covering 3 different types of sports including track events (e.g. 400m), field events (e.g. javelin), teamwork events (e.g. beach volleyball) were observed. It is noted that cameras installed in a stadium tend to focus on one or a group of athletes or on the moving path of the athletes. In addition, objects in these video scenes tend to always stay within particular regions of the video image for each camera. Users may be interested in one particular screen region e.g. the screen centre. Users may also be interested in multiple screen regions (variant preference) associated with multiple objects of interest (invariant preferences) across different live video frames. In addition, the boundaries of regions of interest are dynamic and can change as sports events progress. To support this, a system is required to understand the user zooming regions of interest in terms of both size and number of those regions.

4.5.2 Selective Target Zooming Personalisation Model

Here, personalised selective target zooming is modelled in terms of the interaction during zoom-in.

4.5.2.1 Overview

The proposed sub-system architecture (Figure 4-17) consists of five major components, the user profile, video ZUI control, personalised zooming control, streaming server and terminal. This design is similar to the personalised multi-angle viewing sub-system. This also leverages the advantage of a network centric approach, all components are linked though the video ZUI control. The video ZUI control is integrated into the video player that is networked to a video streaming server. The video ZUI control triggers the personalisation of the zooming task and presents the results of this on the video player screen.

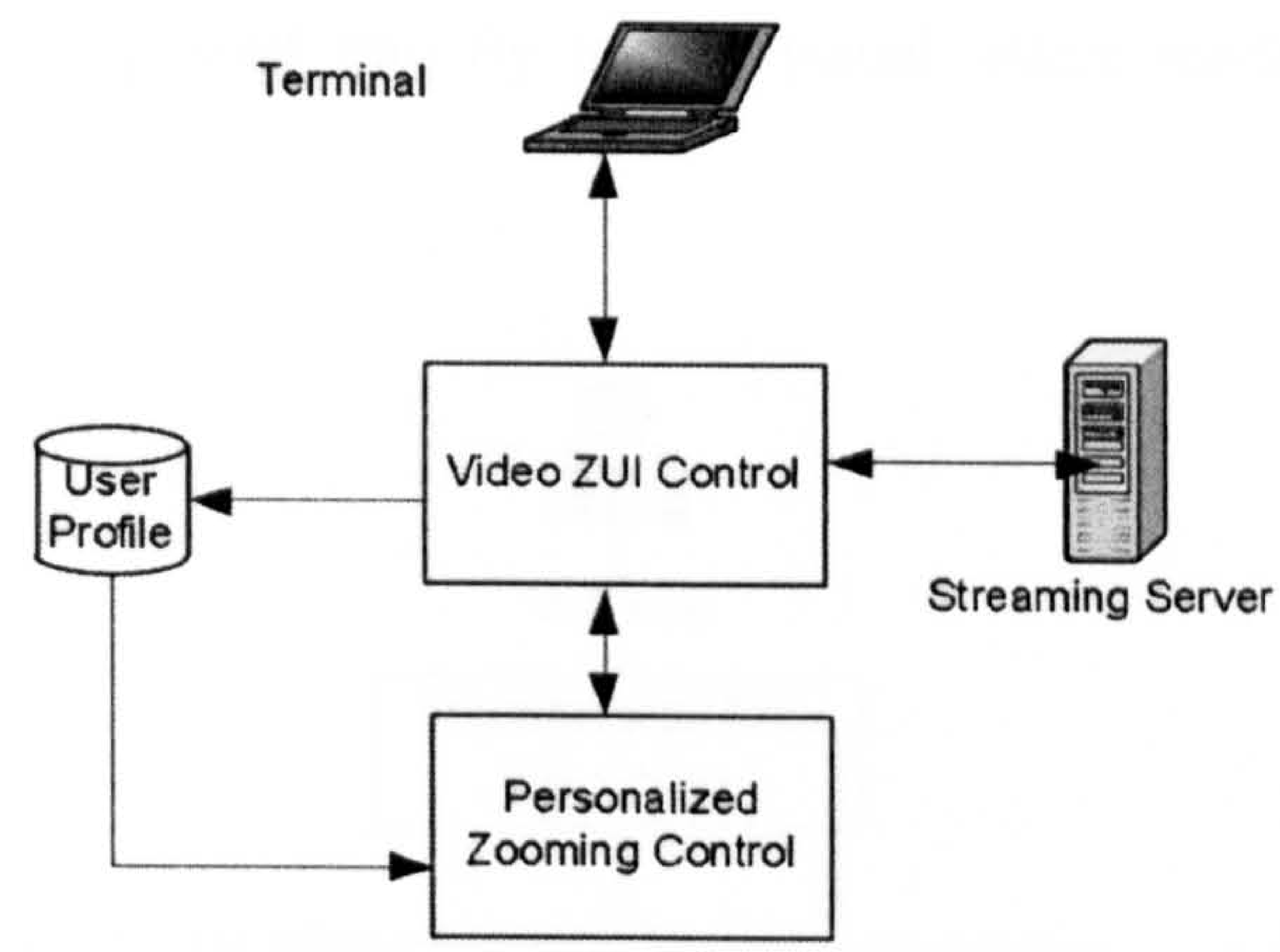


Figure 4-17 System architecture for a personalised ZUI

The personalised zooming control analyses a user's historical interactive data for the zooming task and generates a predicted zooming target position based upon a user's zooming preference. The user profile stores the historical zoom data. Figure 4-18 presents the personalisation model that defines the relationships between components including the personalised zooming control, user interactions, visual effect module (which contains zooming animation), time-shift playback, video quality adaptation and user profiling. Users can enable or disable the personalisation but the visual effect module and user profiling carries on working. Such a design allows the system to implicitly and continuously update a user's profile so that more historical user data can be used in the personalisation process. Each of these components is described in more detail in the next few sections.

4.5.2.2 Task User Interface and User Interaction

The user interaction to support consists of the following sequence, locate the zooming target, start the zoom-in and zoom-out to default state. The last two steps can be automated to support personalised, or active adaptive, zooming. The personalisation decision making process is used to detect and differentiate the cues for non-personalised versus personalised zooming. The user input is defined as either a non-personalization request input or a personalization request input. The non-personalization request input allows the user to go through the following tasks: a) locating the zooming target b) starting the zoom-in c) zoom-out to default state. In a personalization request input scenario, user only needs to perform the last two steps, i.e. b) and c).

When a request for task personalisation is input by the user, the system passes the request to the personalisation zooming control for further processing. A trigger for non-personalised zooming is passed directly to the visual effect module. User input is recorded in the user profile.

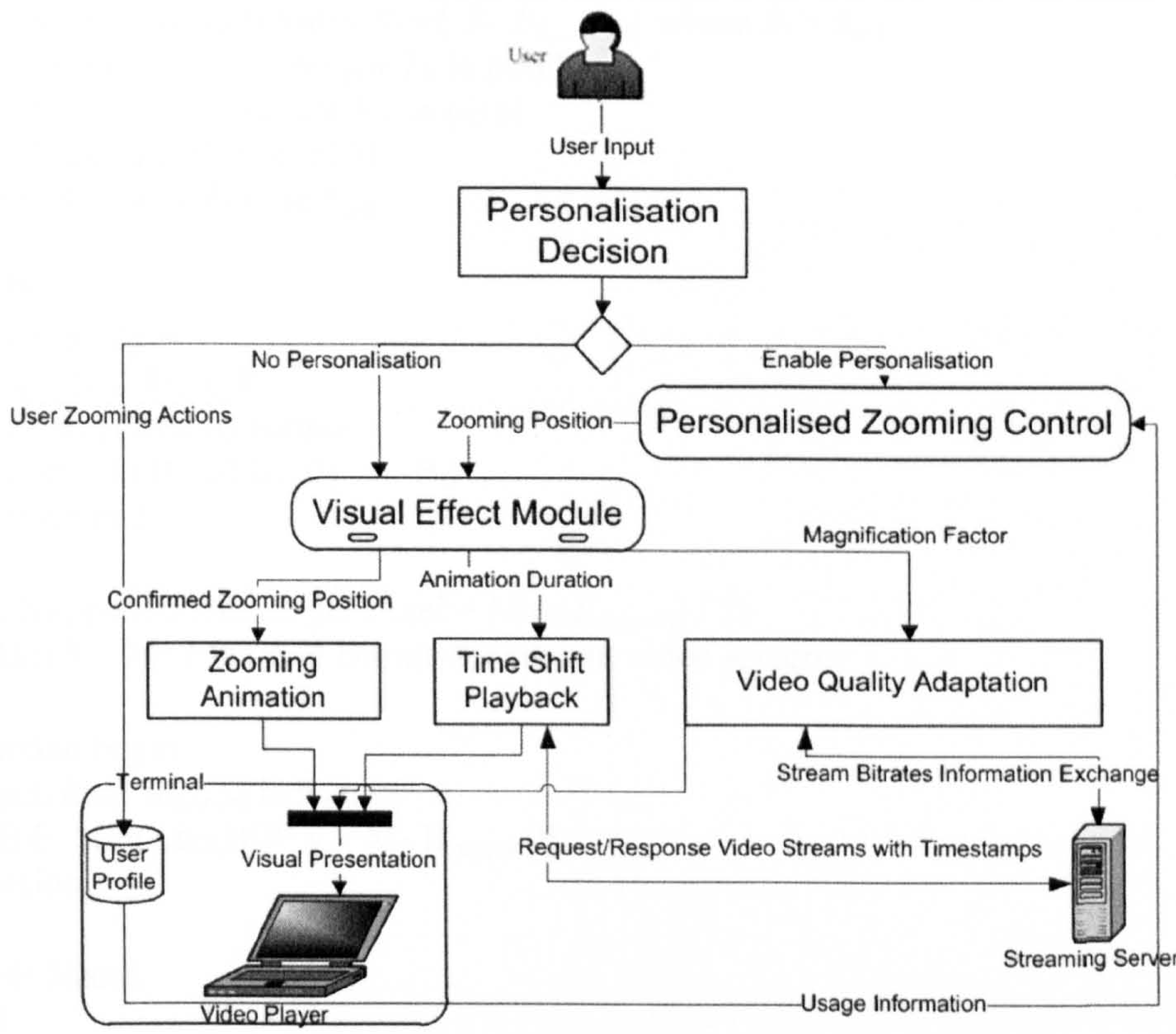


Figure 4-18 Selective target zooming personalisation model

The visual effect module extends the existing ZUI support by introducing three additional sub-processes, video quality adaptation, time-shift playback and zooming animation. Video quality adaptation is video screen resolution driven – it automatically streams video at high bitrates for large screen. The video quality adaptation assumes that the source video content is encoded with different quality levels and adaptively streamed. An algorithm proposed (Table 4-11) extends the algorithm proposed in Table 4-7 to achieve video quality adaptation based upon a set of factors including available bitrates, terminal screen size, current video screen size and magnification level.

A time-shift playback process allows the current video to be playback past video frames before the zoom-in animation process begins. Therefore, the ideal time-shift would be the duration of animation, i.e. if the zoom-in occurs at time T_n then the playback timeline will be at T_n - T_{ani}, where T_{ani} is the animation duration.

Table 4-11 Video quality adaptation algorithm

```
Input: Video stream bitrates Bs={ B1, B2,..., Bn} where Bn> Bn-1,
       Terminal screen height Ts in pixel,
       Video screen height Vs in pixel,
       Magnification level M
Output: Adaptive Bitrate Badp

BEGIN
Recursion begin
for each item Bsi in Bs
If Bi is supported by terminal
Bsupport ← add Bi to{ B1, B2, ..., Bi}
Recursion end

BiP←Supported Bitrate per Pixel = Max(Bsupport) / Ts
Init MaxB= 0← Maximal Bitrate for current video screen;

Recursion begin
for each item Bsupporti in Bsupport
MaxB← Min(Abs(BiP*Vs*M- Bsupporti))
Recursion end

Badp ← MaxB
END
```

4.5.2.3 User Profile Modelling

User profile processing is a process of accessing, parsing and extracting usage information. In the proposed personalised zooming control model, a user profile contains the following usage related information: the normalized coordinates of the zooming target central position, sports event name and camera view name. These data are encoded in XML and used in the active adaptation process.

Table 4-12 Personalised zooming user profile in XML


```
<?xml version="1.0" encoding="utf-8"?>
<!--Created 13:12:10-->
<Profile>
 <PZ id="AI1">
 <user id="UserID">
 <session id="61" cx="32.1089297023433" cy="33.1378299120235" event="M100F1"
camera="High" />
 <session id="35" cx="33.2488917036099" cy="54.8387096774194" event="M100F1"
camera=" High " />
   <session id="453" cx="72.625" cy="60.6741573033708" event="M400F" camera="
High " />
   <session id="499" cx="40.1875" cy="61.0486891385768" event="M400F" camera="
High " />
   <session id="1" cx="52.5" cy="46.9413233458177" event="M400F" camera=" High "
/>
   <session id="582" cx="45.75" cy="72.2846441947566" event="M400F" camera="
High " />
   <session id="602" cx="28.1875" cy="55.9300873907616" event="M400F" camera="
High " />
  </user>
 </PZ>
</Profile>
```


An example of user profile is given in Table 4-12. Each zooming action is recorded as a session item associated with an id. The coordinates of a zooming position is normalized to a value between 0 and 100 in the X and Y axis. The 'event' attribute indicates the viewed event and the 'camera' attribute denotes the labelled camera view. For example, the zooming histories for two events (men's 100m final and men's 400m final) are recorded for high camera views on the roof of the stadium.

The usage information extracted from the parsed user profile normally share the same event and same camera view corresponding to the current viewing event and its associated camera view. The extracted usage information afterwards can be partially or fully fed into a region of interest clustering process (see section 4.5.2.4) as a training data set.

4.5.2.4 Personalised Zooming

Personalised zooming control enables the system to automatically distinguish the regions of interest for zooming and rank them in terms of preferences based upon an analysis of the zooming usage information stored in the user profile. Figure 4-19 shows the personalised zooming state chart which consists of two processes of user profile processing (previous section) and regions of interest clustering (modelled here).

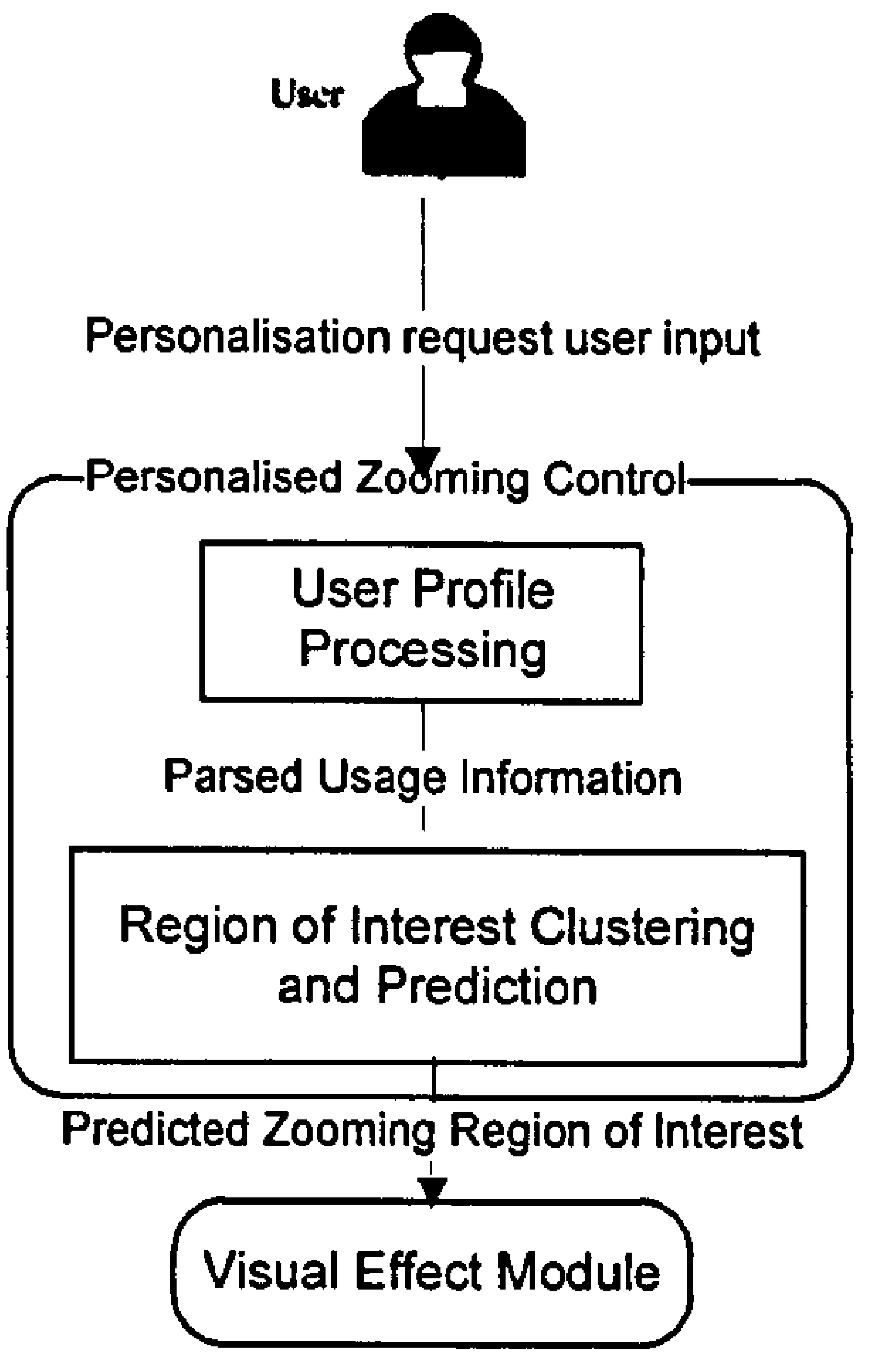


Figure 4-19 Personalised Zooming Control Model

Region of interest Clustering: Clustering identifies groups within a set of unlabelled data. Data partitioned into the same group are similar with respect to a measurement metric, e.g. the Euclidean distance.

There are different types of clustering techniques as surveyed (Höppner et al., 1999). The approach taken here is fuzzy type clustering. The main advantages using fuzzy clustering is that data can belong to more than one cluster; clustering is based upon the strength of membership of that data in each cluster. Data with similar high degree of membership are clustered to the same group.

In the personalised zooming model, clustering users' areas of interest from a particular camera view is challenging. This is because visual objects within a defined display area cannot be located absolutely. In this case, the cluster membership feature of fuzzy clustering can be used for the case where data boundaries are not clearly defined.

The Fuzzy C-means (FCM) algorithm (Bezdek, 1981) is a well-known fuzzy algorithm that allows data to become a partial member of different clusters with a membership value between 0 and 1. FCM divides N data points into C fuzzy groups and identifies the cluster centre in each group via the minimization an objective function (see equation 4.5.2.4-1):

$$J_m(U,V) = \sum_{i=1}^{N}\sum_{j=1}^{C} \mu_{ji}^{m}||X_i - C_j||^2 \qquad 4.5.2.4\text{-}1$$

Where:

U a matrix of (μ_{ji})

V vector of cluster centres $\{Vi\ (i = 1, 2 \ldots c)\}$

μ_{ji} membership of ith data in the jth cluster

X_i feature coordinate of ith data

C_j the jth cluster centre

$||X_i - C_j||^2$ distance of X_i from C_j

m (>1) the degree of fuzzification

c (>2) the total number of clusters

Two critical parameters for FCM are the total number of clusters and the initial cluster centres which are normally problem domain determined. For the region of interest clustering process, these parameters can be determined by the algorithm shown in Table 4-13. The number of clusters is determined by the magnification level, e.g. a 1.3 magnification (i.e. enlarge 1.3 times) could produce 3 clusters. The initial clustering centres are obtained by a counter clockwise rotation of the farthest point about the centroid of zooming points coordinates taken from the user profile.

Table 4-13 Personalised zooming control cluster algorithm

```
Input: Zooming focal points coordinates Zc = { Z1, Z2,..., Zn }
       Magnification level M>1, Total number of clusters Nc
Output: initial cluster coordinate Ci = { C1, C2 ,..., Cn }
        Total number of clusters Nc

BEGIN
 initialize cnum =0;
 initialize zcentroid =0;
cnum← round((M-1)*10);
 zcentroid←Point(Mean(∑Zcx), Mean(∑Zcy))

intialize maxdis = 0;
intialize maxpoint =null;
Recursion begin
  for each point Zi in Zc
  if distance(Zi , zcentroid) > maxdis
update maxdis;
maxpoint ← Zi
Recursion end

initialize CP =null
Recursion begin
  repeat Nc times
 CP ← Add{
```

$$C_n = \begin{pmatrix} zcentroid_x \\ zcentroid_y \end{pmatrix}^{-1} \begin{pmatrix} \cos\partial & -\sin\partial \\ \sin\partial & \cos\partial \end{pmatrix} \begin{pmatrix} zcentroid_x \\ zcentroid_y \end{pmatrix} \begin{pmatrix} maxpoint_x \\ maxpoint_y \end{pmatrix} \}$$

```
Recursion end

   Nc ← cnum
   Ci ← CP
END
```

A clustering process can be conducted after the initial clustering centres are determined for all the clusters. Table 4-14 describes FCM clustering steps with extracted zooming focus points from a user profile.

Table 4-14 FCM clustering algorithm

Step 1: Initialize c fuzzy cluster centres C_i based upon the extracted zooming coordinates and calculate the membership degrees of recorded zooming centres where the following conditions are satisfied: a) $\mu_{ji} \in (0,1)$, $1<i<N$, $1<j<C$ b) $\sum_{i=1}^{C} \mu_{ji}^{m} = 1, 1 < i < N$ c) $0 < \sum_{i=1}^{N} \mu_{ji}^{m} < N, 1 < j < C$
Step 2: Calculate the next values for cluster centres: $V_j = \frac{\sum_{i=1}^{N} \mu_{ji}{}^{m} x_i}{\sum_{i=1}^{N} \mu_{ji}{}^{m}}, j = number\ of\ clusters, i.e. 1, 2, ..c$
Step3: Update the fuzzy degree of membership: $\mu_{ji} = \frac{1}{\sum_{k=1}^{C} (\frac{\|\|x_i - c_j\|\|}{\|\|x_i - c_k\|\|})^{\frac{2}{m-1}}}$
Step 4: If the currently calculated values *Vi* for the cluster centres are not different from the values calculated at the previous step (subject to error e), then stop.

Zooming Region of Interest Prediction: It is envisioned that the clustered areas will be more stable with minor boundary changes when the zooming times increase. In this thesis, a null hypothesis based heuristic function is used to obtain a possible future zooming region of interest. The null hypothesis is that there exists a median with a highest p value for the historical chosen zooming region indices (two-tail test). The steps to determine the future zooming region of interest is summarized in Table 4-15.

Table 4-15 Future zooming region of interest determination algorithm

Step 1: propose hypothesis *H0*: there exists a median with highest p value of the historical chosen zooming region of interest indices
Step 2: Choose significance level p
Step 3: Use indices medians as the test statistic
Step 4: Determine the accept region for the statistic
Step 5: Get the index with the highest p

In some situations, a user may change their zooming preferences, e.g. a user becomes more interested in athletes. This may be due to several factors such as a significant video content change, e.g. an incident happened in the spectator area or zooming action mistakes occurred. The consequence of this change will directly reduce the future zooming region of interest prediction precision. In order to mitigate this noise data, the amount of training data (i.e. recorded zooming centres) used for zooming region of interest clustering is adjusted according to the prediction rate changes. The amount of training data decreases accordingly to allow the system predicted zooming preference to align to a user's recent zooming preferences (see system implementation section Chapter 6).

4.6 Time-Shift Viewing

4.6.1 Time-Shift Viewing User Model

Based on the user model proposed in section 4.2.2, time-shift viewing user model is presented in Figure 4-20. Time-shift viewing allows a user to view previous scenes of interests from a list of highlighted scenes. In order to achieve this, the scenes should firstly be highlighted and user's preference of scenes should be defined. Because interaction is heavily live content driven and because the interaction to select video highlights itself requires some directing expertise which is not possessed by most normal users, this is challenging to automate. To simplify the problem, user preferences for scenes to highlight are related to preferences to view sports incidents. Hence, sports incident identification is needed.

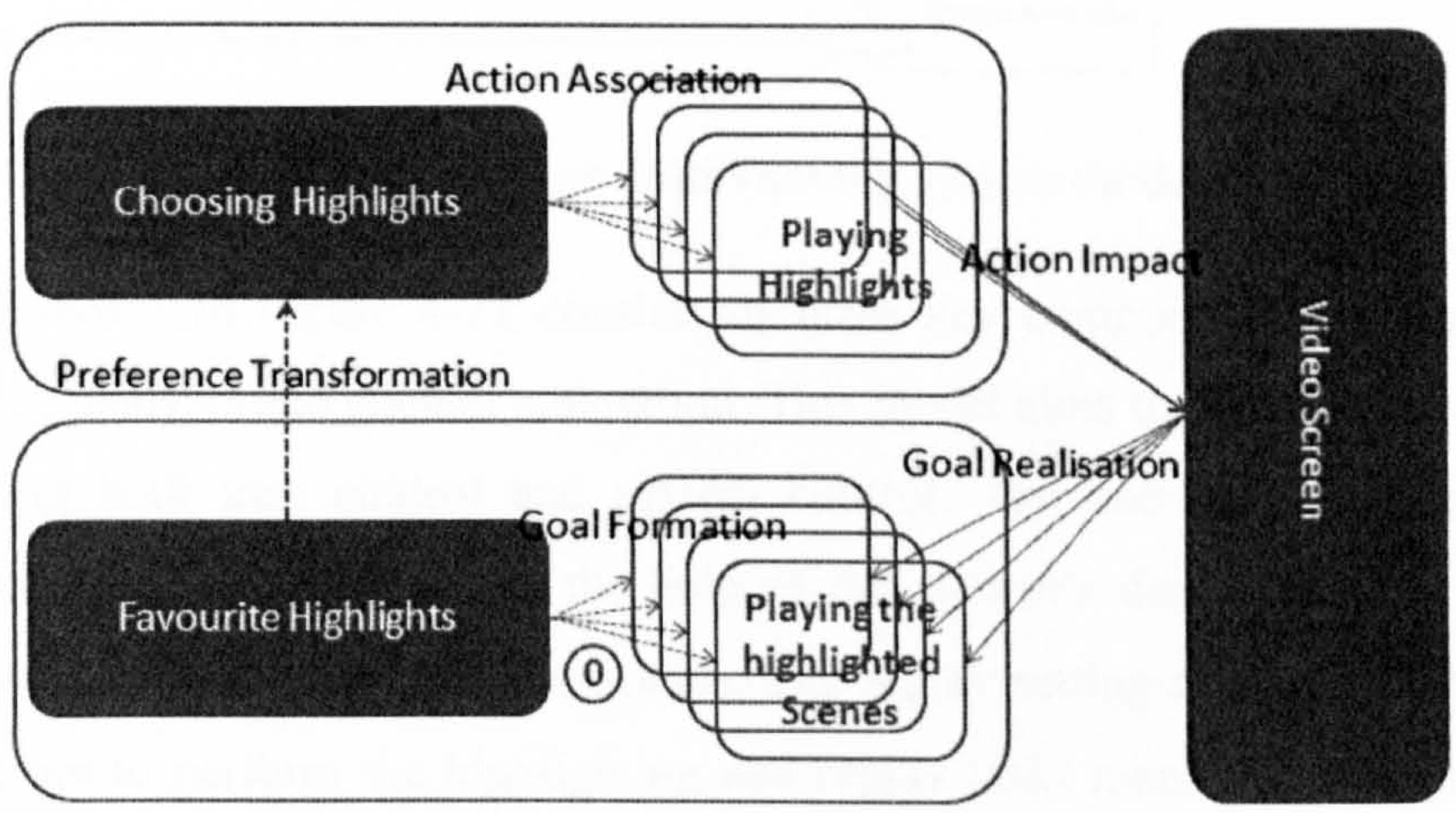


Figure 4-20 Time-shift viewing user model

4.6.2 Time-Shift Viewing System Model

Time-shift viewing is modelled in terms of incident-highlight and auto-replay interaction.

4.6.2.1 Overview

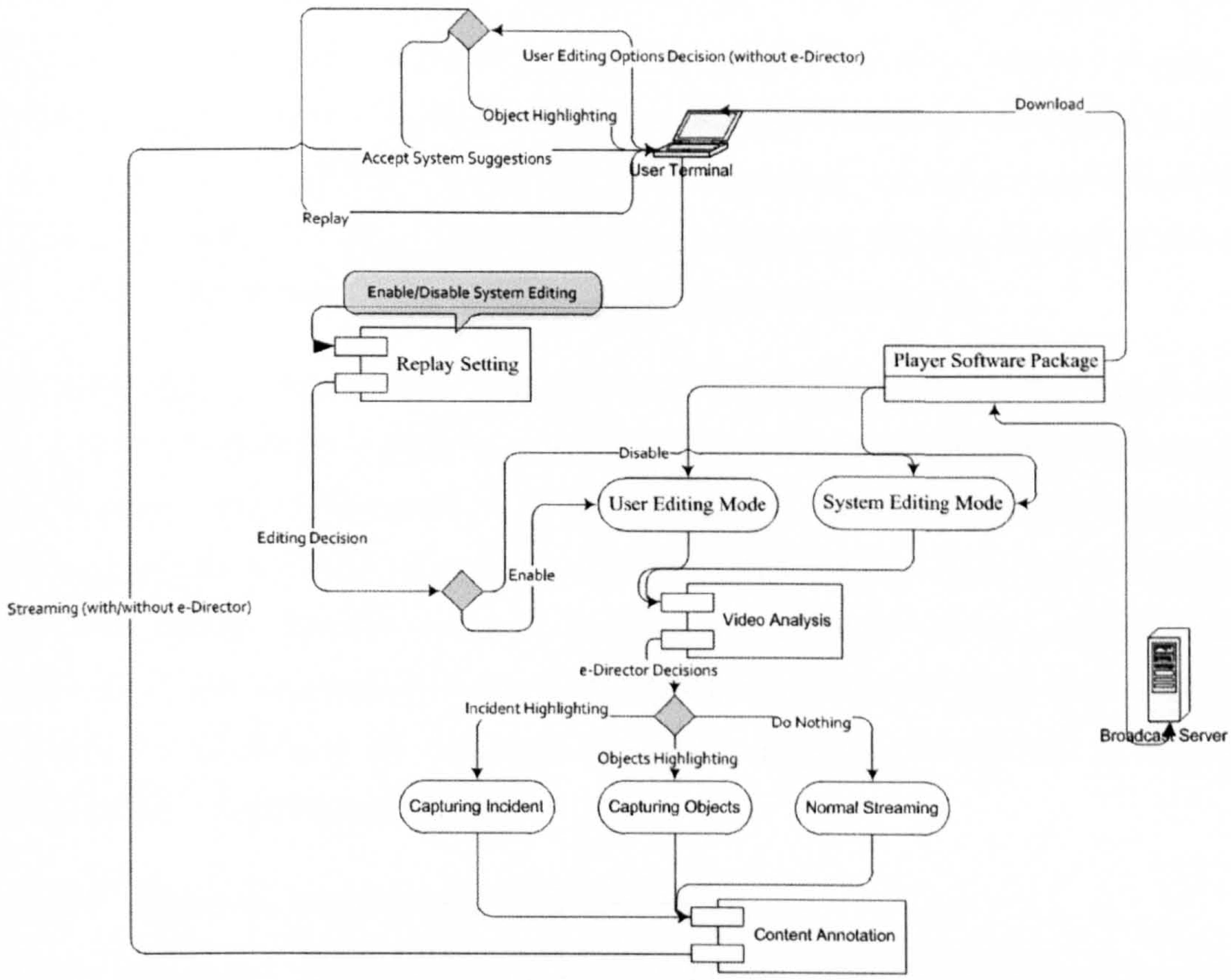


Figure 4-21 Time-shift viewing system model

The model shown in Figure 4-21 consists of three key components including, replay setting, video analysis and content annotation. This model aims to leverage the advantage of combining both user control and system control. The user has the flexibility to manually highlight the scenes with the help of the system's decisions or to allow the system to automatically highlight the scenes. The replay setting allows a user to decide whether or not to perform the highlighting and replay tasks manually or automatically. The video analysis component analyses the video content and the content annotation component annotates the video content respectively. In the following sections, these components are modelled in more detail and related the (Web-based) user interface requirements for interactions.

4.6.2.2 Director Expertise

This requirement gives the overall logic for the video analysis used to simulate a director's ability to highlight parts of a video to automate time-shift viewing. The director rules used to support this here mainly address three specific director tasks, incident highlighting, objects highlighting and auto-replay. In video production professional handbooks such as (Millerson and Gerald, 2001) (Holland, 2001) (Millerson, 1999), contend that replay can be summarized as a balance between a dynamic and compelling sequence of images. One implication of this balance is that some images in a sequence of images are more compelling than others, sports incidents are typical examples of this. Sports incidents can either be expected athletes actions or unexpected ones that normally, strongly, impact on the event progress.

In sports events, a slow change in image content between successive video frames may indicate that something is going to occur, e.g. a kick, ready to run at start line, whereas a quick image content change may indicate something is happening and may impact the game outcomes, e.g. a goal scored, crossing the finish line (Perry et al., 2009). Although this work did not explicitly point out in detail how a director should highlight sports video sequences, a general principle is that quick image content change may require a review of what actually has happened. In contrast, additional details may be required when there are small image changes.

4.6.2.3 Task User Interface and User Interaction

This requirement is achieved via the replay setting described in Figure 4-21. The live editing task interface triggers an automated director and informs users about this. The interactive task determines who is going to direct the sports events (system or user) whereas the system tasks are more concerned about the execution of a director's action. Table 4-16 lists the editing tasks for both system and user respectively.

Table 4-16 Interaction and live editing task interface

Interaction	Interface / System
Highlight Scenes (image frames and objects)	Play, save or delete the highlights, Visualise / disable the annotation
Reminder to replay highlighted scenes	Enable/disable replay

4.6.2.4 Active Director

The image motions are the data used in the adaptation process. The active (system) director models video analysis and content annotation (Figure 4-21). It identifies the image motion conditions. It executes a director's time-shift actions such as highlight incidents, highlight objects and auto replay.

Image Motion Condition Detection: Motion detection techniques are frequently used to detect changes between video (image) frames. In this model, video motion condition detection is based upon pixel motion analysis. The proposed model, assuming that image motion changes have a Gaussian distribution in most circumstances, uses a statistical model to classify image motion.

A pseudo Gaussian distribution, given a certain statistical significance level, for an example image motion change is shown in Figure 4-22. The inside of the critical area is defined as an unbalanced area. More specifically, the right tail critical area indicates fast image motion whereas the left tail critical area indicates slow image motion. Motion changes can be described from two perspectives. One is the motion difference (MD) between two separate video frames, and the other is the difference between the MDs. The former indicates the changing conditions of an image in a whole sequence of images. The latter indicates the motion change conditions in a whole sequence of images.

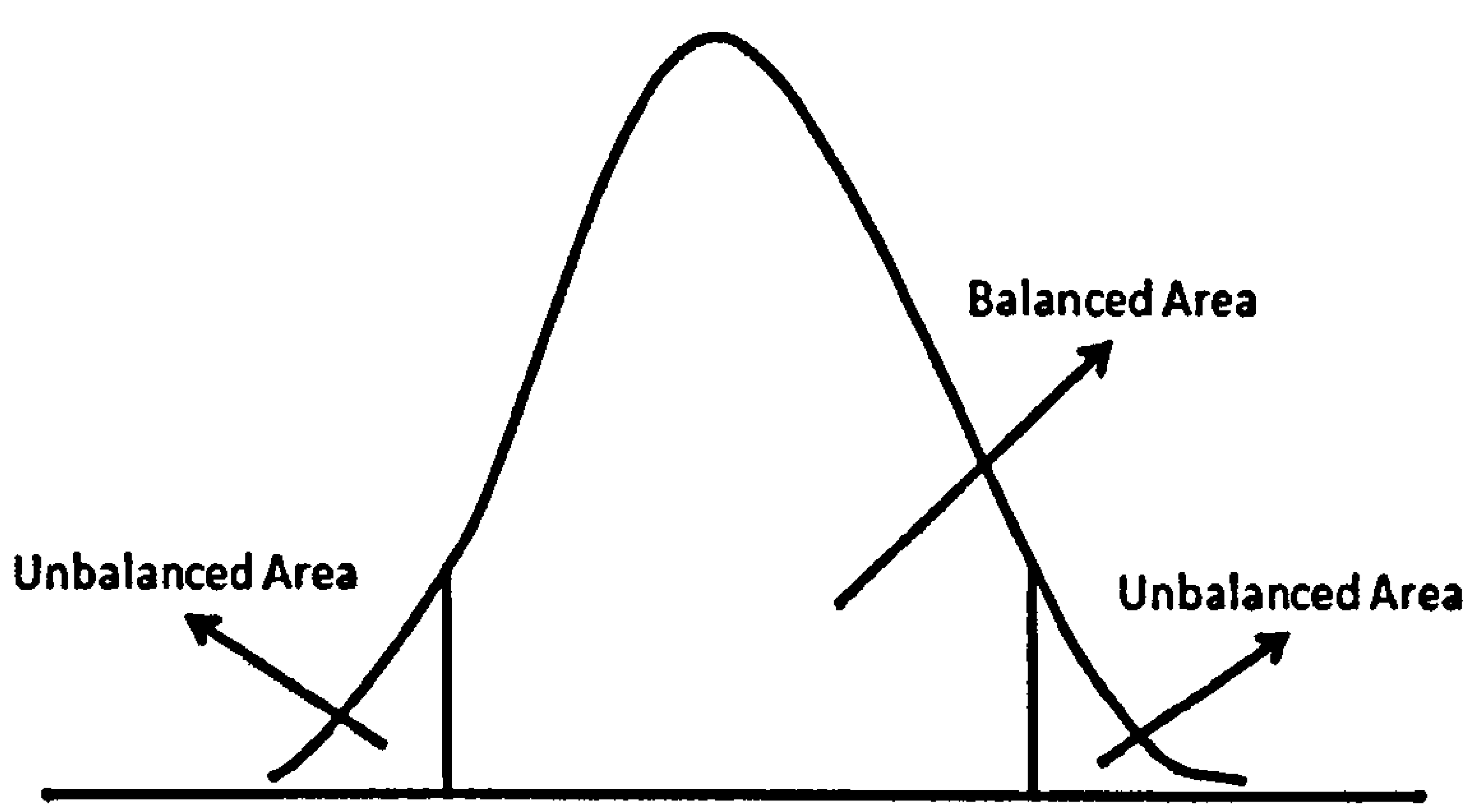


Figure 4-22 Motion change distribution with respect to a Gaussian distribution

Figure 4-23 summarizes motion change conditions with their inhabited regions (i.e. either left tail or right tail) in Figure 4-22. The matrix shows four different unbalanced motion conditions between images and between sequences of images.

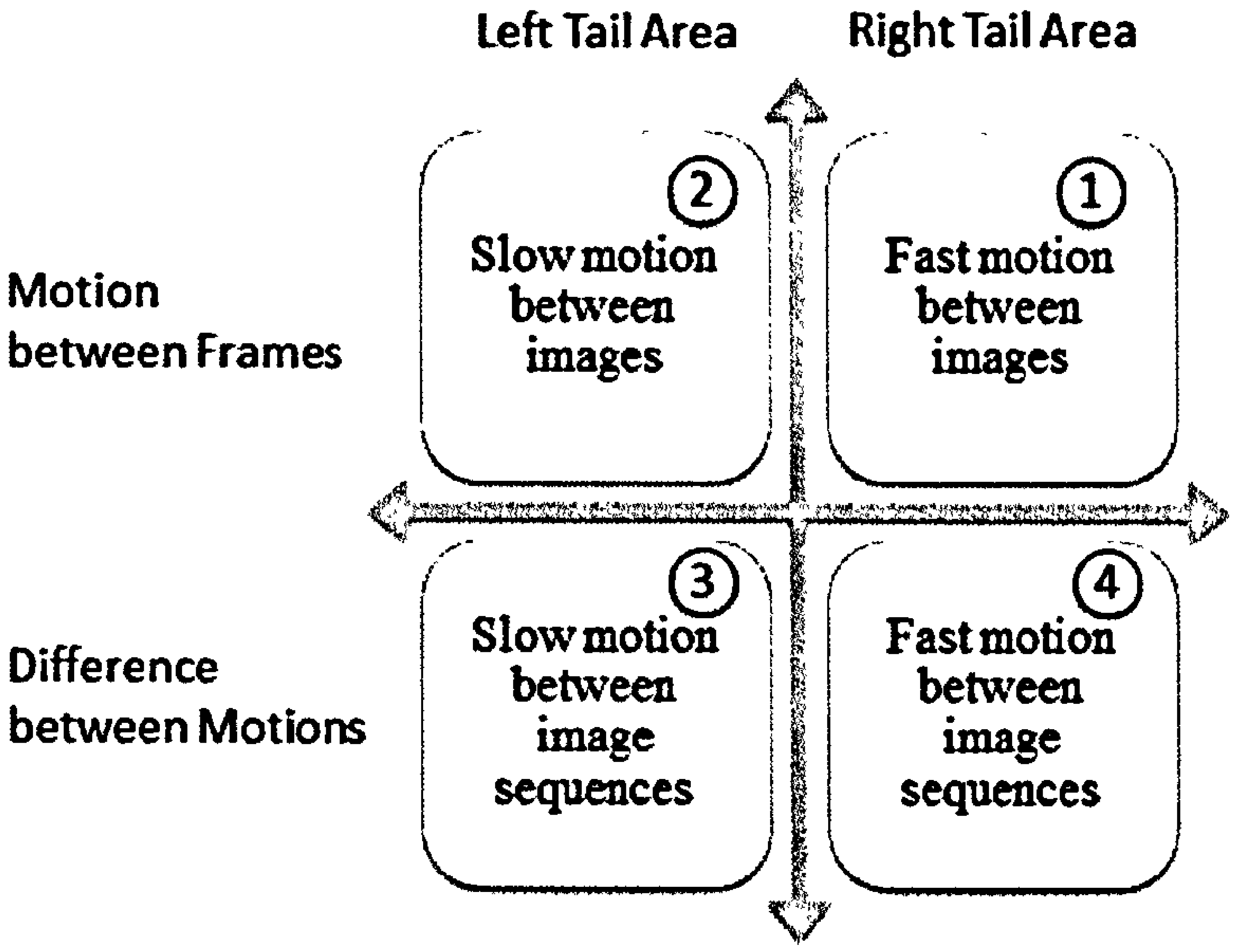


Figure 4-23 Motion change matrix

Here, condition 1: fast motion between images and condition 3: slow motion between image sequences (Figure 4-23) are chosen as the director options trigger conditions. Incident highlighting in live sports events is more concerned with a longer incident prediction lead time, i.e. the longer the amount of time before an incident of interest happens, the better. Therefore identifying MD between frames can be a better approach than detecting differences between MDs (condition 1) as the former has half the minimum triggering time for the latter. For example, if given an fps = 25, the former approach has a minimum triggering time of 25/f second and the latter approach has a minimum triggering time of 50/f second, where f is the image sampling frequency. In contrast, object highlighting, requires the system to find a moderate temporal interval to draw a user's attention to a highlighted object otherwise a user may not have enough time to notice the highlight. In this case, detecting the difference between MDs could help achieve this (condition 3). Following the previous example, given fps = 25, the motion detection between image sequences requires a minimum interval of $50/f$ second and the counterpart approach requires a minimum interval of $25/f$ second. Likewise, the *auto replay* can be triggered when condition 3 is active. This is because the replay can be an effective tool to enrich video when there is relatively less content per unit time. Among four conditions in Figure 4-23, condition 3 indicates a longer period of time with less content load, i.e. a minimum period of $50/f$ second according to the example given.

Incident Highlighting: According to the director rule given earlier, incidents are associated with fast motion images. The incident highlighting model is based upon a high level image processing technique.

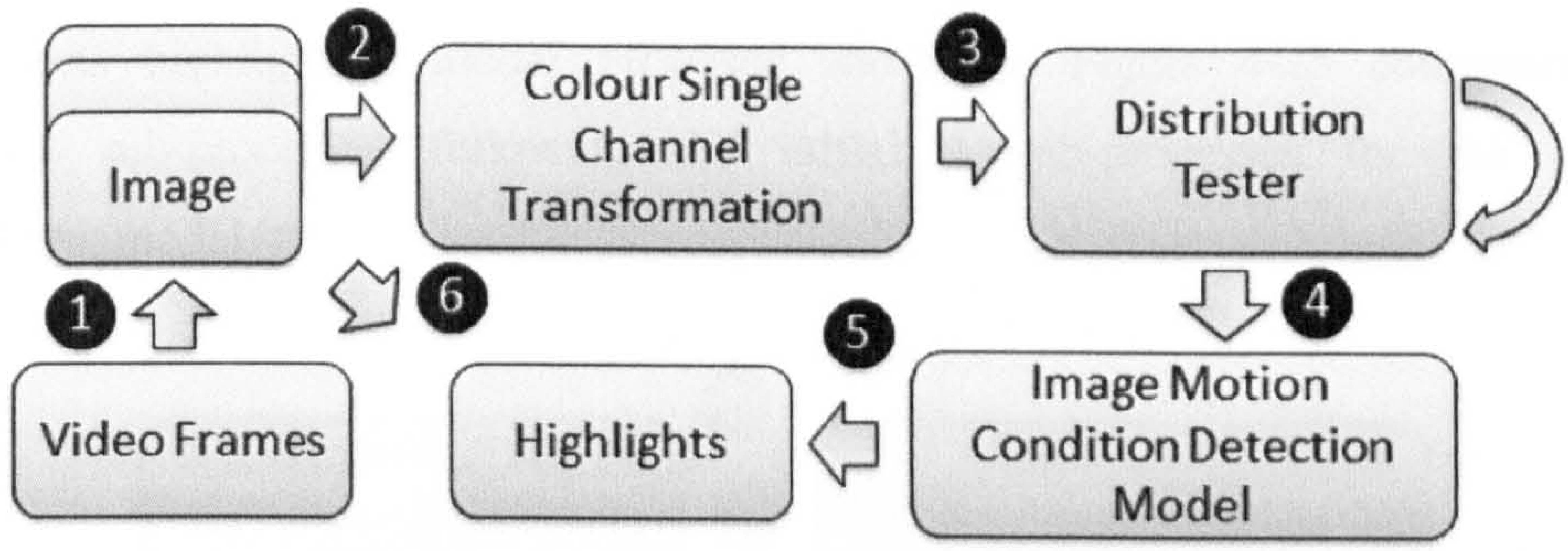


Figure 4-24 Highlight incidents model

In Figure 4-24, the proposed incidents highlighting model consists of six sub-processes that represent the life cycle of the highlight incidents process.

1. *Image extraction*: key frames are captured from a video stream
2. *Colour channel scaling*: ARGB channels are converted to single grey scale channel so that each pixel can be presented with a single value to facilitate the comparison between images.
3. *Gaussian distribution test*: the set of motion change values $(I_{j,k} - I'_{j,k})$ are tested The Gaussian property, $I'_{j,k}$ denotes an image motion change at time T. $I_{j,k}$ denotes image motion change at time $T+1/f$ of a pixel in jth row and kth column. f is the sampling frequency. The test will be iterated unless the collection is Gaussian distributed.
4. *Motion condition detection*: determines whether or not the motion change value at T time matches condition 1 (see Figure 4-23) in a Gaussian distribution by using a Z score test.
5. *Highlights annotation (timestamp)*: the timestamp of the captured frame is used to annotate the highlights.
6. *Highlights annotation (image)*: the key frame image is captured as the cue for highlights.

Object Highlighting: Highlighting objects in video content enriches the viewing experience and the content. The highlight objects director's option proposed here is focused more on the latter. According to the director rule proposed, video content that matches condition 3 (see Figure 4-23) tends to have a less content load. Thus highlighting objects in a sequence of images gives viewers a visual cue about what objects in the content they focus on. For example, during a goal kick in a football match, the goal keeper can be highlighted.

The objects highlighting model proposed shown in Figure 4-25 comprises three additional sub-processes following the initial 4 sub-processes for the incident highlighting model.

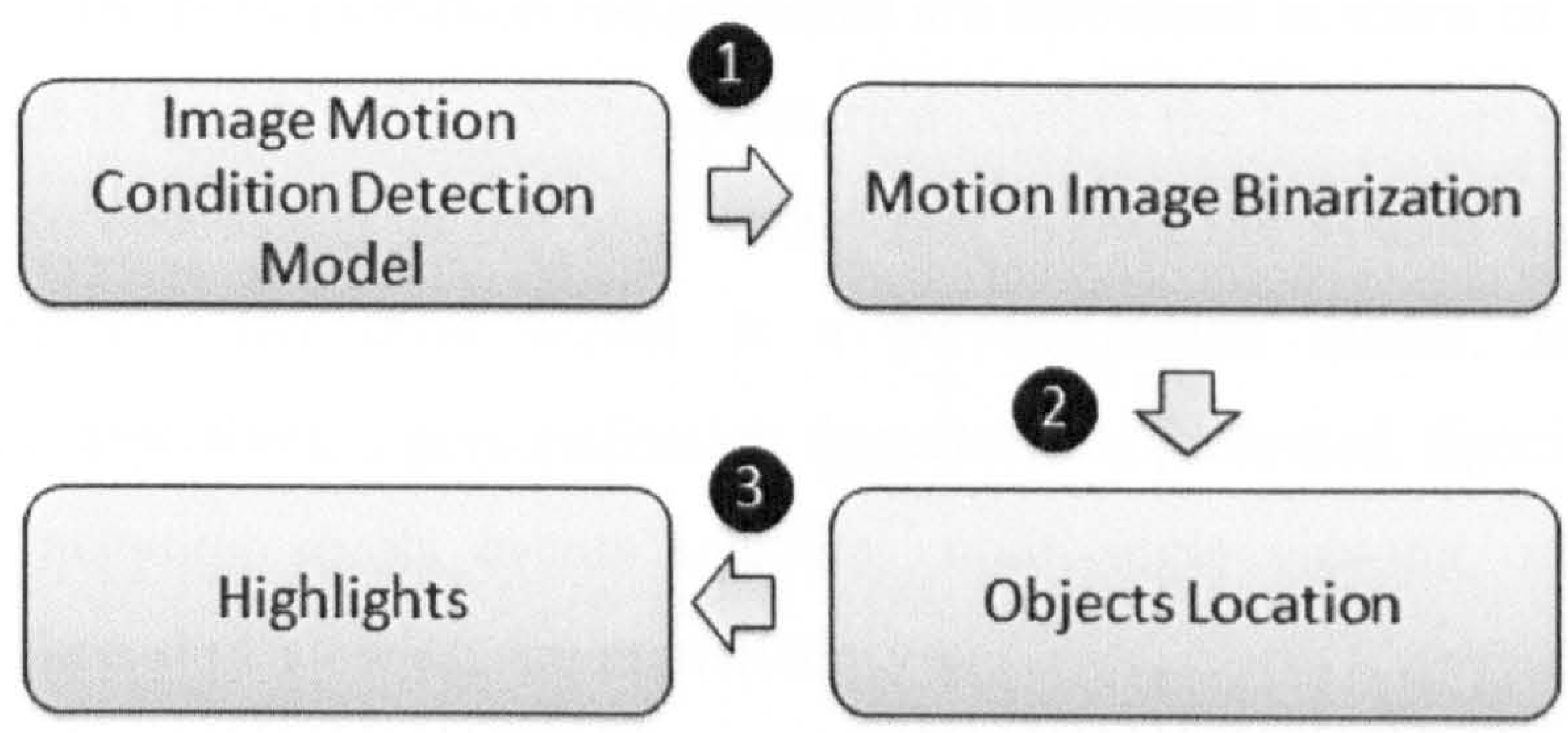


Figure 4-25 Object highlighting model

- *Binarize motion image*: the image describing the motion between two frames is bianrized, e.g. the motion part is white and the static part is black. The purpose of this is to facilitate the location of objects
- *Locate objects*: the object with a significant motion is located in the image. The highlighting centroid $Centroid(x, y)$ has the following properties based upon the image histogram analysis:
    a. $x_c = \boldsymbol{Max(Histogram_{horizontal})}$
    b. $y_c = \boldsymbol{Max\ (Histogram_{vertical})}$

The boundary (i.e. $\boldsymbol{X_{MinandMax}\ and\ Y_{MinandMax}}$) of the highlighted objects has an area that has the following properties:

$$\boldsymbol{Histogram\ of\ x_{MinandMax}} > \frac{\boldsymbol{Max(Histogram_{horizontal})}}{\boldsymbol{2}}$$

$$\boldsymbol{and}\ x_{minandmax} - x_c = \boldsymbol{Min}(x_{minandmax} - x_c)$$

$$Histogram\ of\ y_{MinandMax} > \frac{Max(Histogram_{vertical})}{2}$$

$$and\ y_{minandmax} - y_c = Min(y_{minandmax} - y_c)$$

- *Highlights Annotation*: the process of visually annotating the located object such as a visual overlay.

4.7 Summary

In this chapter, the personalisation requirements are addressed in terms of three aspects. The first aspect is the objectives of the interactions within the live sports events viewing system. The second aspect is the user model that allows the system understands the user's preferences. The third aspect is a personalisation model. In support of personalisation objectives, a personalisation framework is presented. Specific task-based interactions, including sports events selection, multi-angle viewing, selective target zooming and time-shift viewing, are modelled.

Chapter 5

5 Evaluating Personalisation Performance in a Next Generation A-V Player

This chapter addresses the challenges in evaluating personalised interactive systems. Hypothesis Testing based evaluation forms the basis of the approach. This will be extended to address three practical issues. They are: What is to be evaluated? (system performance versus user experience); How are the evaluation results acquired? (discrete versus continuous evaluation); Who is/are involved in the evaluation? (single versus multiple users).

5.1 Overview

Only about a quarter of studies evaluated a personalisation system after personalisation took place (Van Velsen et al., 2008), i.e., personal profiles are usually gathered before the onset of the first user session. Studies do not seem to evaluate a personalisation system when the personalisation is taking place (Soui et al.'s work, 2008). In this thesis, personalisation is evaluated principally at run-time at the operational stage. Usability also includes coarse-feedback about some aspects of personalisation is evaluated in a post-session questionnaire (Section 6.3). The (personalisation) evaluation model is composed of a series of statistical and hypothesis based tests that are used in concert with the core concept of the personalisation evaluation, i.e. the functional parameters defined in each of the two stages (Figure 5-2). The model diagram is shown in Figure 5-1.

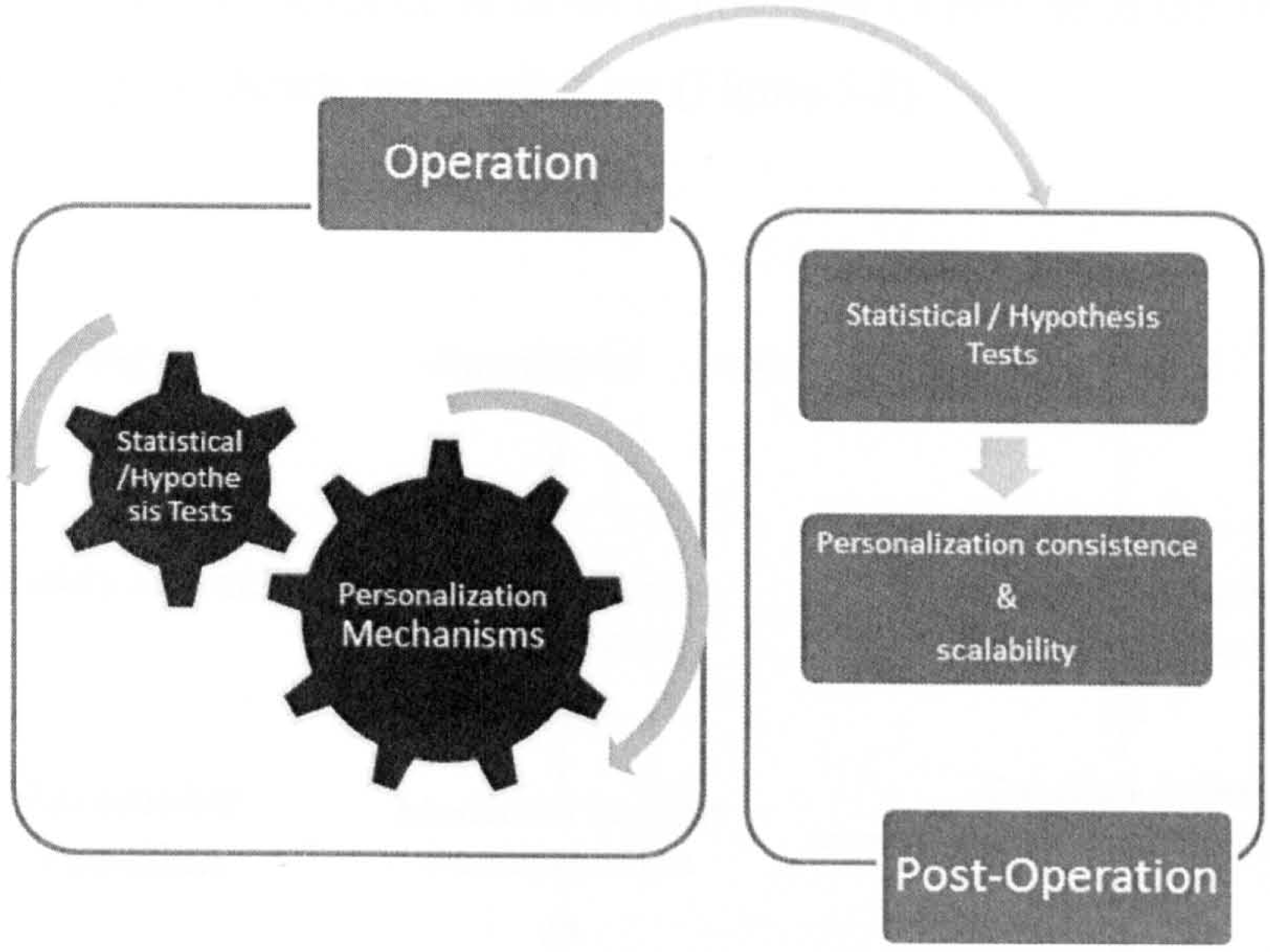


Figure 5-1 Personalisation evaluation model

At the operational stage or run-time, personalisation mechanisms can be validated, e.g. how well the system learned from users. This also supports fine tuning of the personalisation mechanisms at runtime. The statistical and hypothesis testing methodology used is designed to be independent of particular personalisation mechanisms.

The tests at the personalisation post-operation stage are designed to justify the overall system in terms of two aspects of the personalisation performance, namely *personalisation consistency and scalability*. Personalisation consistency indicates the similarity of personalisation effectiveness across different users as specifying by a particular metric, e.g. prediction precision. The scalability of a network, system, or process is defined as the ability to maintain or improve its performance at light, moderate, or heavy input loads (Bondi, 2000; Jogalekar et al., 1997).

5.2 Test Parameter Acquisition

Test parameters for a personalisation system are metrics that reflect the quality of the personalisation system. For Hypothesis Testing, the test parameter must be statistically measurable and its meaning most relate to a quality of the personalisation system. Both prior knowledge and qualitative research tools can be used as a supporting tool to define such parameters. However, the drawbacks surveyed in section 2.4 of using prior knowledge and qualitative research tools may prohibit the evaluation process. Here, the sources of parameters are identified in terms of how they contribute to the verification of different quality aspects during personalisation (Figure 5-2).

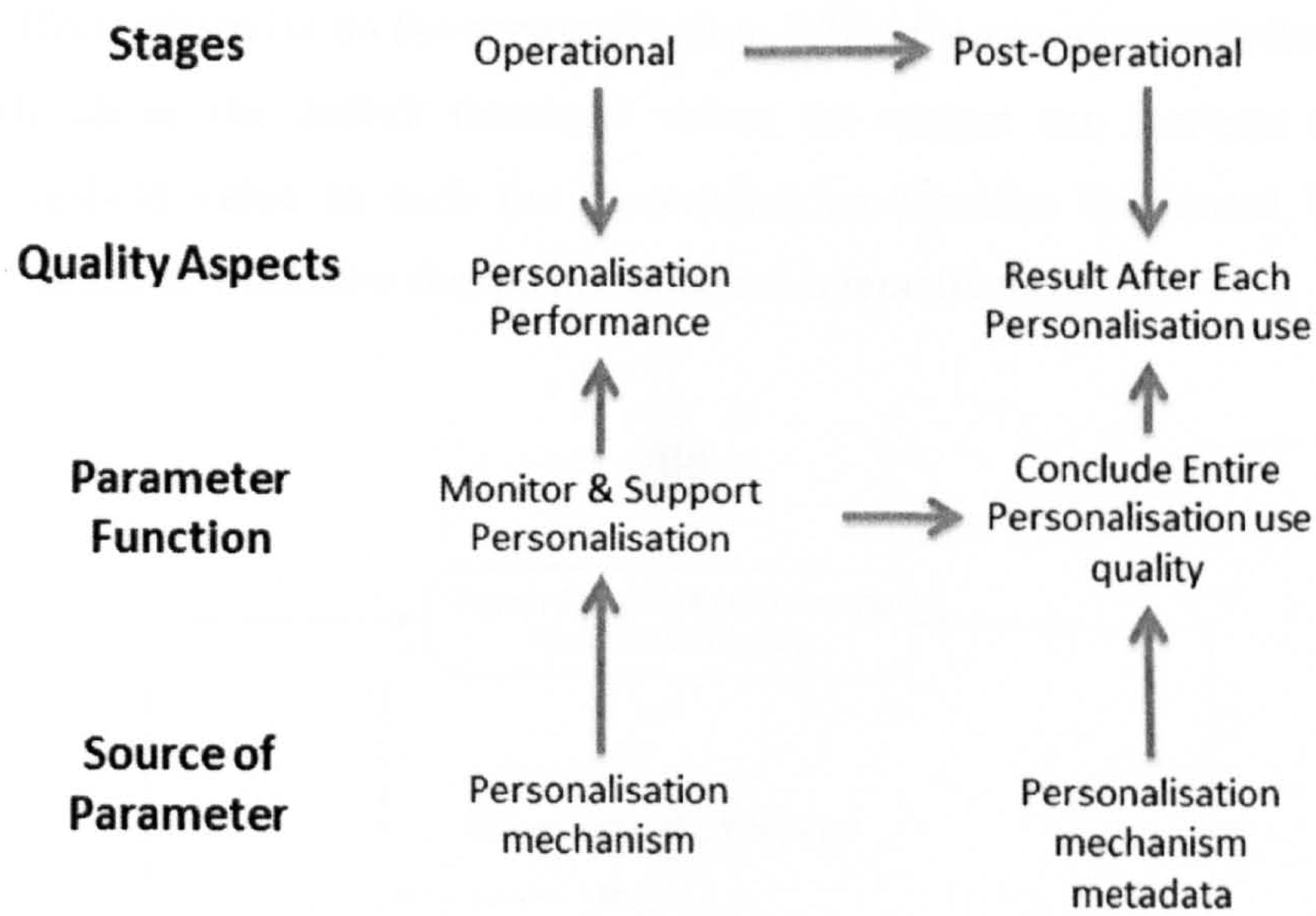


Figure 5-2 Personalisation quality aspects and sources of evaluation parameters

At the operational stage, the test parameters can be derived from the underlying personalisation mechanisms, e.g. a parameter can determine how well the system's model of the user preferences matches a user's actual preferences. These test parameters are used for two purposes, one is to monitor the personalisation implementation performance, and the other is to ensure the personalisation is useful and usable at runtime. Testing parameters connote the overall performance of personalisation in terms of system requirements such as scalability, consistency etc. These parameters, to an extent, act as summative data that describes the overall performance of personalisation mechanisms. They can also be used as metadata that describes the accumulative values of the testing parameters used in operational stage, e.g. the mean of each use session's prediction precision.

In the following sections, detailed models are dedicated to evaluations at both operational and post-operation stages.

5.3 Personalisation Evaluation at Run-Time

The core of the evaluation algorithm will be a Hypothesis Test that assesses the test parameters. A test parameter can be obtained from the target personalisation mechanism. Based upon the hypothesis test this parameter, the personalisation function can be performed or fine-tuned. Here is an example, given the prediction precision on recommendation as a test parameter, a default threshold value can be used to determine whether the system should perform the personalisation based upon a simple rule such as if the average prediction precision is below this value then no personalisation should be taken into effect, otherwise do the personalisation. When the average prediction precision is constantly above the default threshold value, the system can increase the default precision threshold value as such the personalisation function fine-tuned. Figure 5-3 shows the evaluation workflow for the personalised interactive system.

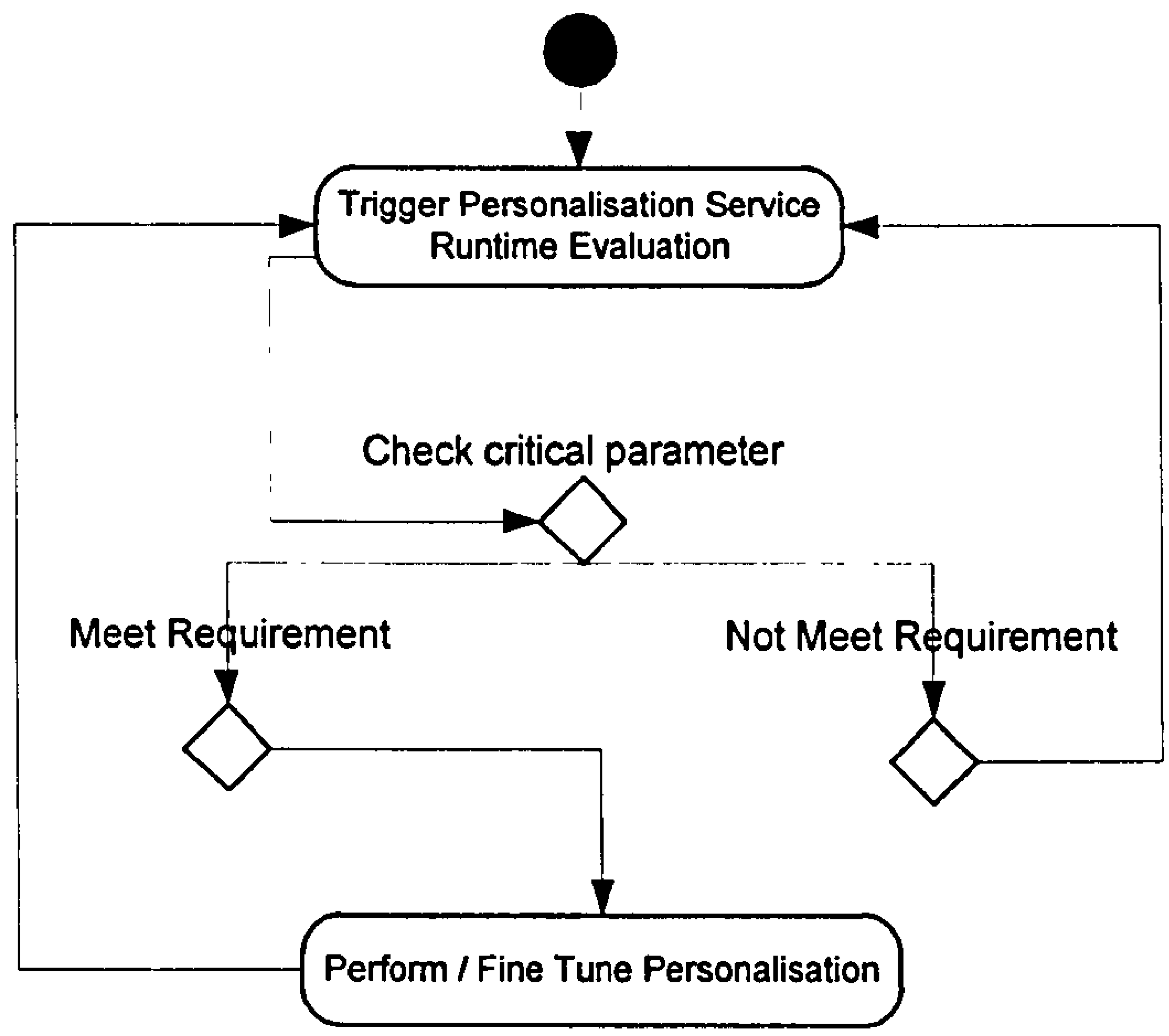


Figure 5-3 Runtime personalisation evaluation workflow

Personalisation evaluation also needs to be conducted after each system use session. The motivation of doing such evaluation for a personalisation system is to test personalisation consistency and scalability. Evaluation during system operation promotes service quality at runtime and evaluates the effects of personalisation on multiple users and service consistency and scalability issues in terms of increased quantities of user input and user number of the system.

5.4 Consistency Test

The consistency test (i.e. similarity of the personalisation effectiveness across different users) algorithm consists of a set of tests as shown in Figure 5-4. The whole algorithm is triggered by a consistency test request from a system administrator. Each test sample consists of a parameter with a set of values and is retrieved from an individual user when personalisation is being used. The samples with normal distributions are put into a similarity test which will tell whether these users have an equal variance in a particular test parameter or not. The samples with an equal variance for a particular parameter e are noted as positive samples. A test statistic α, e.g. a median, can be obtained afterwards from the samples with normal distributions. Samples classified as non-normally distributed are then respectively tested against α with a hypothesis e.g. non-normally distributed samples have a higher median than α. If the hypothesis is proven true, the sample will also be marked as a positive sample. Finally, the *consistency rate* can be obtained, i.e. number of positive samples divided by the total number of samples.

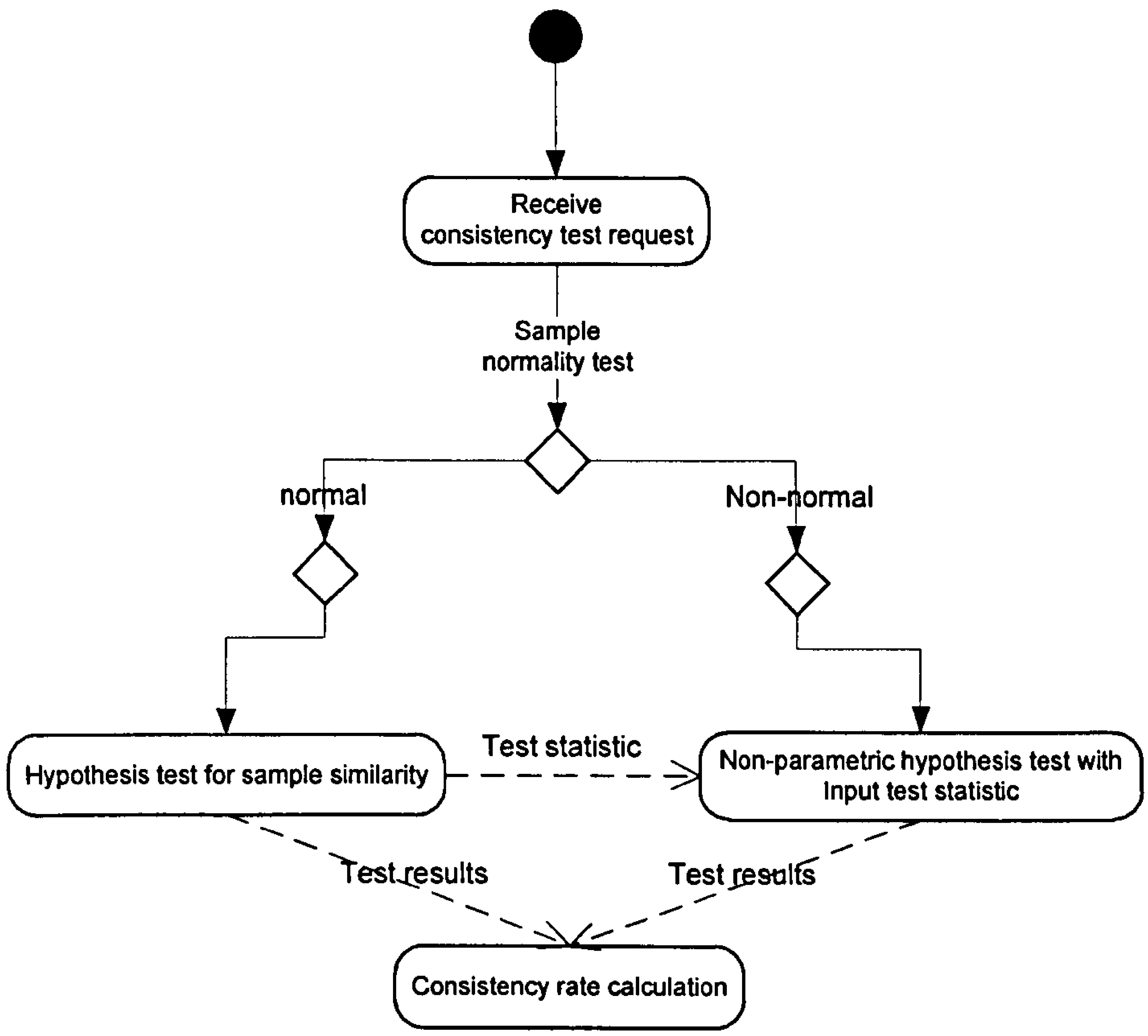


Figure 5-4 Personalisation consistency evaluation algorithm workflow

Statistical methods such as Jarque-Bera Test, Shapiro–Wilk test, Kolmogorov–Smirnov test and etc. can be selected to test if a sample is normal. A similarity test can utilise the F-Test based methods such as one-way ANOVA (Johnson et al., 1995). The non-parametric test for non-normal samples can use statistical methods such as Wilcoxon signed-rank test (Wilcoxon, 1945)

5.5 Scalability Test

The scalability issue in this thesis is viewed in terms of the use of both the back end and front end sub-systems. The back end, the focus is on the quantity of users and how this can affect personalisation quality. At the front end, the focus is on the number of times each individual user actually personalises a task and how the increasing number of uses can affect personalisation quality (see Figure 5-5). The choice of test parameter to reflect personalisation quality will be dependent upon the personalisation mechanism to be evaluated, e.g. in the sports events selection interaction, the recommendation accuracy is chosen (see section 5.6). The core test method is a correlation test that assesses a linear relationship (i.e. no relationship/ascending relationship/descending relationship) between a test parameter and variable pair. Here, the 'number of users' is used for multi-user scenarios such as group based recommendations. The number of times an individual user uses personalisation, e.g., for selective target zooming, is recorded.

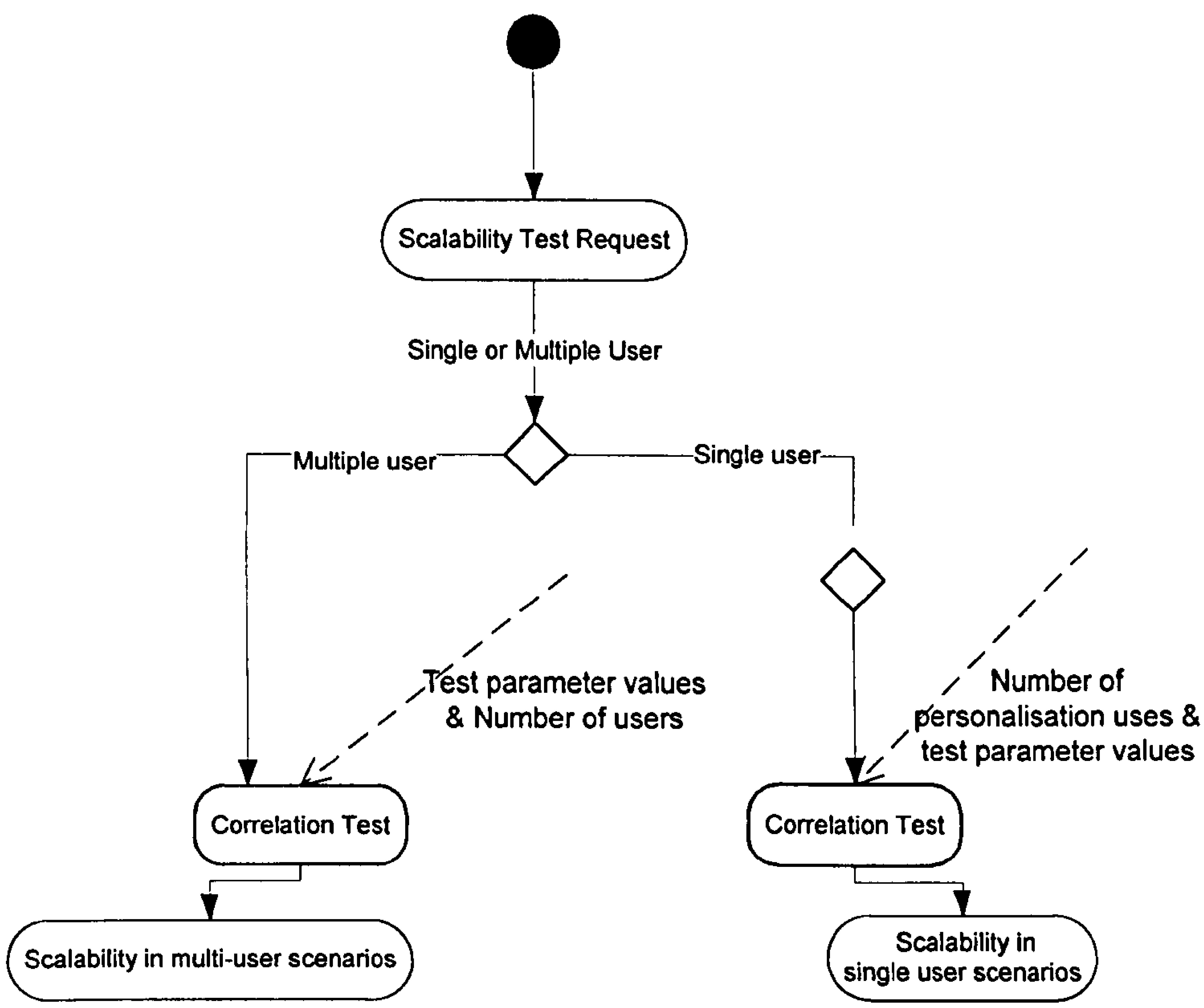


Figure 5-5 Personalisation scalability evaluation algorithm workflow

Among existing methods for testing correlation, the Pearson correlation coefficient (Rodgers and Nicewander, 1988) is well recognized for testing normally distributed samples. When the sample is not certain to be normally distributed, a Spearman's rank correlation (Myers and Well, 2003) can be used to get more accurate results.

5.6 Passive and Active Personalised Interaction Evaluation

5.6.1 Passive System Adaptation

Personalising sports events selection is mainly achieved via by a passive approach, namely a recommender system. The effectiveness of personalisation depends on a user's decision (for individual based recommendation) or a group's decisions (for group based recommendation) for recommended items, i.e. whether they have chosen the recommended top-k sports events. In order to evaluate the personalisation, a test parameter representing both users' past selections and the recommended top-k events is required. Here, the recommendation accuracy is expressed as:

$$\varepsilon = 1 - \frac{Position_i}{k} \text{ (recommendation to individuals)} \quad (5.6.1\text{-}1)$$

$$\varepsilon = \frac{1}{n}\sum_{j=1}^{n} 1 - \frac{Position_{ji}}{k} \text{ (recommendation to groups)} \quad (5.6.1\text{-}2)$$

Where n denotes the number of recommended users, $Position_{ji}$ denotes the position of the jth user selected sports i in the recommended sports list.

Two hypothesis tests can be proposed with respect to recommendations with reference to Figure 5-3, a threshold (median) value is defined to classify the recommendation accuracy as being effective. As a result:

- *H0 (recommendation accuracy):* if the current recommendation accuracy ε_c is greater or equal to a given threshold (median) value, then personalisation is classified to be effective.
- *H1 (with or without recommendations):* if the current recommendation accuracy ε_c is greater or equal to the system without recommendation (i.e. with a random sports events list) then personalisation is effective.

For the consistency test and scalability test, the accumulated recommendation accuracy can be used as the test parameter. For the consistency test, selected test users will be evaluated. For the scalability test, both single users and multiple-users are evaluated in (see Figure 5-5).

The hypothesis tested for the consistency test for each user is:

- *H2 (for recommendation to individuals):* if multiple users have similar recommendation accuracy change pattern, then personalisation is consistent across users.

The hypotheses tested for the scalability test for single and multiple users are:

- *H3 (for recommendation to individuals):* if the recommendation accuracy ε does not tend to decrease, then the personalisation is scalable.
- *H4 (for recommendation to groups):* if the recommendation accuracy ε does not tend to decrease as user number increases, then the personalisation is scalable.

5.6.2 Active System Adaptation

5.6.2.1 Generic Tasks

Personalised interaction for active system adaptation depends on the system's prediction of users' preferences. A prediction precision is (PP) is used as the test parameter and defined as follows:

$$Prediction\ Precision\ (PP) = \frac{Number\ of\ true\ predictions}{Number\ of\ true\ predictions + Number\ of\ false\ predictions} \qquad (5.6.2\text{-}1)$$

The hypotheses tested for the predictability test are:

- *H5 (prediction accuracy):* if the current PP_c is greater or equal than the given median, then personalisation is effective.
- *H6 (compare to non-personalisation):* if the current recommendation accuracy PP_c is greater or equal than the median of previous non personalisation PP_{np}, $PP_{np} = \{PP_{np0}, PP_{np1}, \ldots, PP_{npn}\}$, then personalisation is effective. A random prediction accuracy is used as the non-personalisation prediction accuracy.

Due to the nature of interaction during viewing the sports events, the consistency test and scalability test are mainly targeted at individual users. Hence, the Hypothesis Testing for consistency can be:

- *H7 (for individual users):* if multiple users have similar PP change tendency, then the personalisation is consistent across users.

For the scalability test, the hypothesis can be:

- *H8 (for individual users):* if the PP does not tend to decrease across user sessions, then the personalisation is scalable.

5.6.2.2 Task-Specific Adaptation

Personalisation of this type of time-shift can be sports incident driven, i.e. video content driven. The validation of personalised interaction such as incident highlighting in the simple case considered here is for example, pre-identified timestamps of sports incidents rather for ad hoc incident driven incidents or user detected and triggered incidents. Example task-specific test parameters to evaluate personalisation here are the recall for highlighted incidents and the precision of highlighting in terms of the incident lead time:

$$Incident\ Recall = \frac{Number\ of\ higlighted\ Incidents}{Number\ of\ Predefined\ Incidents} \tag{5.6.2-2}$$

$$Precision_T = \frac{Number\ of\ higlighted\ Incidents\ T\ unit\ time\ before\ incidents}{Number\ of\ Highlighted\ Incidents} \tag{5.6.2-3}$$

5.7 Summary

In this chapter, a personalisation evaluation model is proposed based upon a Hypothesis Testing based approach. The model evaluates a system's personalisation performance at the operational stage and post-operation stage. Different evaluation criteria are used to promote the overall personalisation service quality. At the operational stage, the recommendation accuracy, prediction precision etc. are important. At the post-operation stage, the consistency and scalability of personalisation based upon both individual users and group users are the focus. The evaluation model can be applied to systems to both passive and active adaptation.

Chapter 6

6 Next Generation A-V Player Implementation and Evaluation

In this chapter, the implementation and evaluation of the proposed personalisation models are presented. Three types of evaluation were conducted. One type of evaluation (without users) tests the performance of the system to support key system tasks, e.g., video stream bitrate adaptation. The other type of evaluation assesses the performance of the personalisation. Finally, the usability of the system is evaluated based upon post-use feedback.

6.1 Personalisation Implementation

6.1.1 Overview

Personalisation is invoked using both system interfaces (SI) and user interfaces (UI). The system interfaces support personalisation whereas user interfaces allow user to interact with the system. The system implemented is essentially a customised A-V player (also referred to as the terminal application). It is implemented in C# .NET 3.5 framework compatible class libraries for the Microsoft Silverlight Web application platform. The customisation consists of supporting each of the four main personalised user tasks, and providing the middleware services to support these, e.g., store and retrieve the implicit and explicit user profile. The isolated storage facility of the Silverlight framework is used to store user profiles on the terminal. A Web server (2.5GHZ 64-bit processer and 8GB of system memory) is deployed for both live video content streaming and experimental data acquisition.

6.1.2 Personalised Sports Events Selection (Individual Recommendation)

The system interfaces support live sports feed receiver, user profile data management and user preference processor.

- The live feeds receiver receives smooth streaming video streams and the live sports events information from the video streaming server. It parses the information of sports events name, sports events preference factors and expected length of the events.

- Individual user profile data management maintains this data on the terminal side.

- The user preference processor is embedded in the terminal application which implements the logic of personalisation.

The UI for this sports events browsing uses a horizontal strip menu layout that holds a set of thumbnails representing the sports events in a left to right workflow layout. In Figure 6-1, UI for events browsing is illustrated.

Figure 6-1 Events browsing UI

6.1.3 Personalised Sports Event Selection (Group Recommendations)

The implementation of the personalised sports events selection and play is targeted at groups of users and is similar to implementation of the personalised sports events browsing for individuals. The differences are mainly in the system interfaces that:

- A user profile data store maintains the user data on the server side.
- The user preference processor is hosted on the server side.

6.1.4 Personalised Multi-Angle Viewing

The task interface defined in section 4.4.2.2 is implemented with a class library called Multi-stream adaptation control (for the adaptation and bitrate allocation algorithms).

- *Multi-Stream Adaptation control* receives the smooth streaming streams and returns the corresponding bitrates for them.

The camera switching personalization control implementation (see Figure 4-13) consist of a *Switching interval fuzzification control* (for fuzzification) and *Probabilistic prediction control* (for HMM algorithm).

- The *Switching interval fuzzification control* receives the input read by the user profile processing control and fuzzifies the extracted switching interval values with respect to a particular event. The triangular function used in this thesis defines: a short duration that starts at 0 and ends at mean interval values; a medium duration that starts from the 2/3 of the mean interval value and ends at 5/3 of the mean interval value; the long duration starts at 5/3 of the mean interval value and ends at the maximum interval value.

- The *Probabilistic prediction control* is implemented using a HMM algorithm class library using the algorithm proposed earlier (Table 4-10). This algorithm accepts three parameters including a collection of interval types with corresponding camera types, the previous switching sequence and a posterior sequence of camera types. The result produced is the maximum posterior probability of a camera type with an associated switching interval type. The predicted interval value is the mean of that interval type.

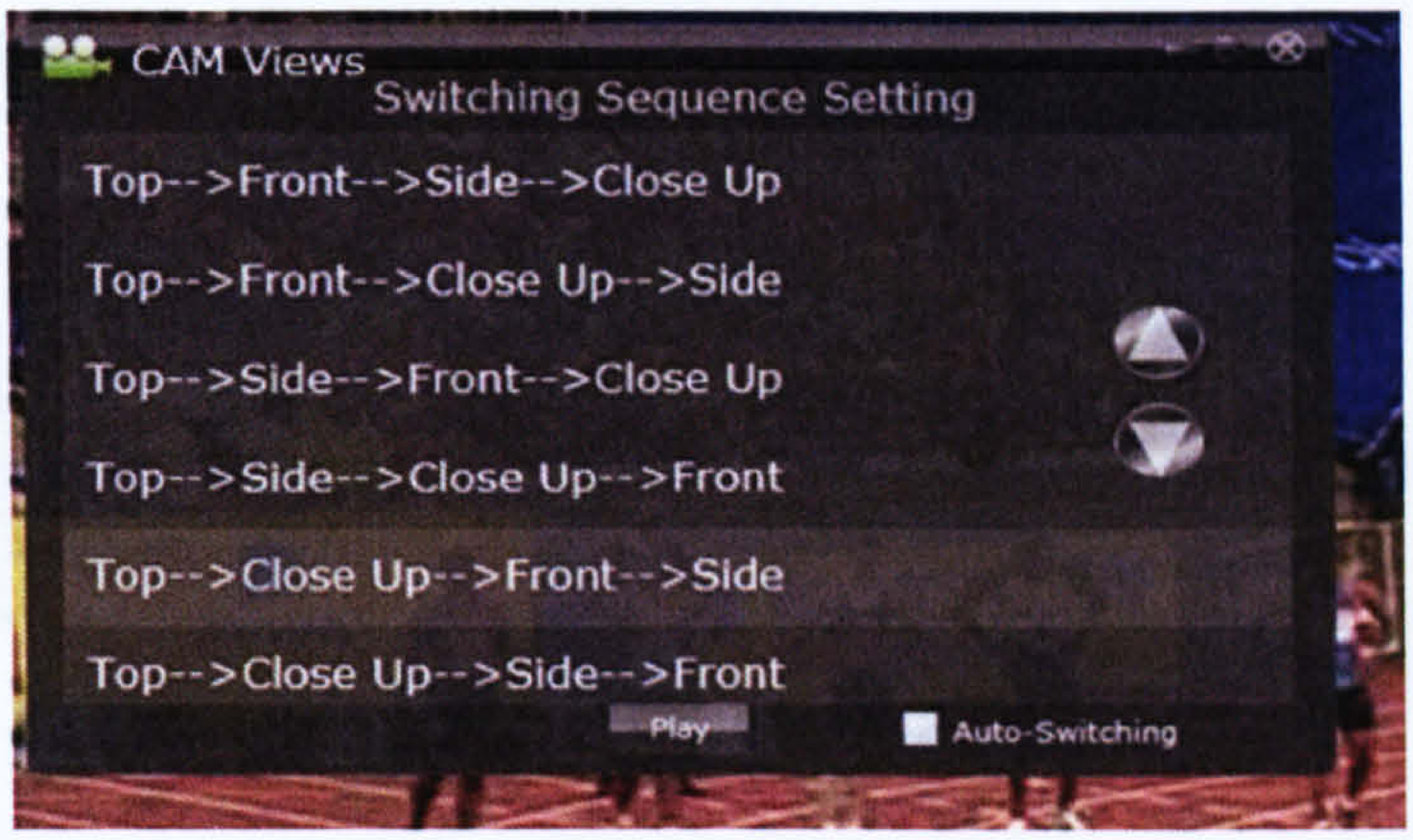


Figure 6-2 Multi-angle viewing UI – camera selection

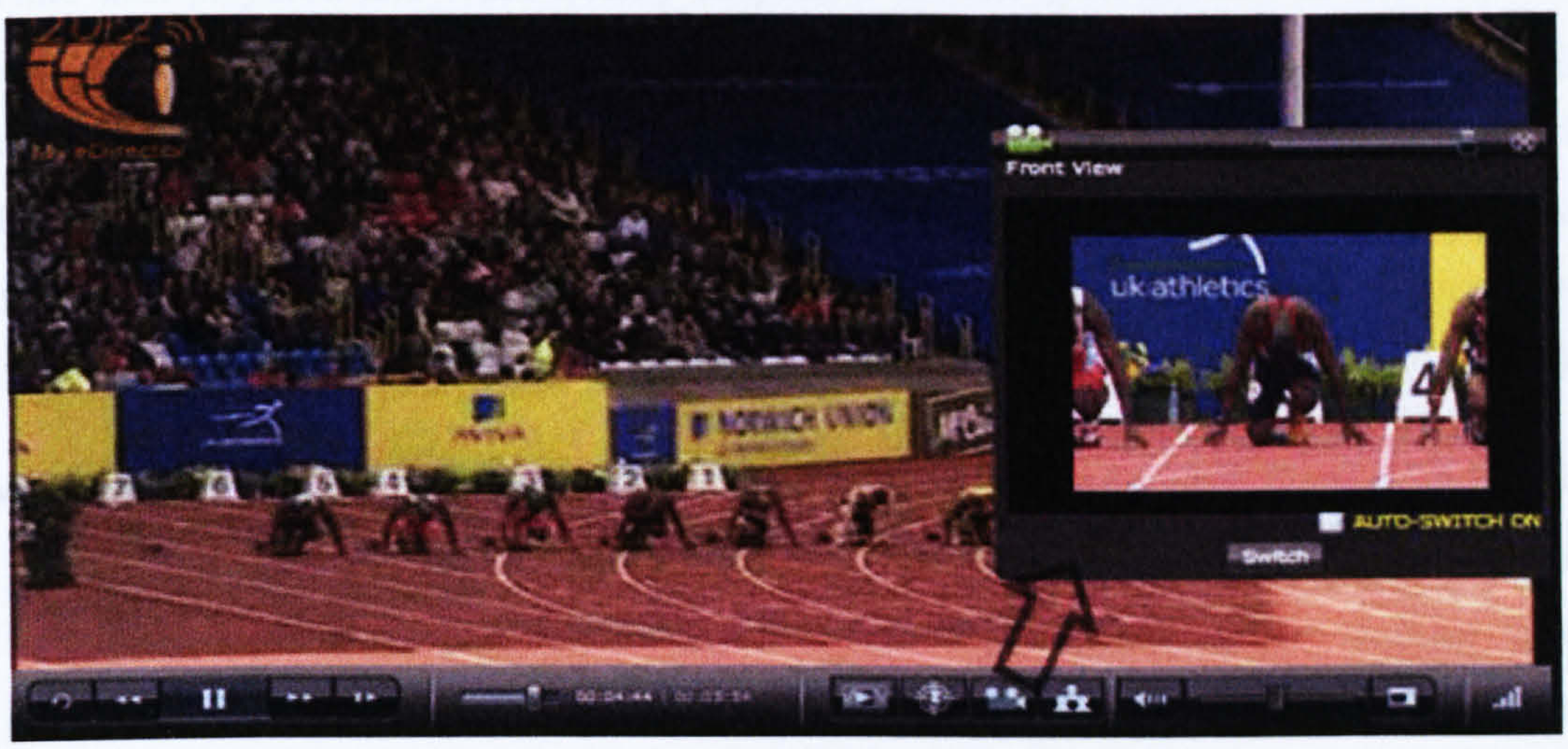


Figure 6-3 Multi-angle viewing UI - camera switching

The camera selection UI (Figure 6-2) allows users to select the cameras of interest in advance. It allows users to manually switch between the available cameras via the ‘switch’ button. An enable / disable auto-switching option is supported via an ‘auto-switching' check box. When enabled, the system can automatically switch cameras based upon user switching preferences that users can submit via a UI form.

6.1.5 Personalised Selective Target Zooming

For the video ZUI, a temporal separated zooming highlight effect is produced by a pixel shader, implemented using a High Level Shader Language[6] (HLSL). The zooming rate can range from 1.0 to 2.0, i.e. 1 to 10 times the zoom of the original video. Zooming implements zooming starting from the original video screen centre to the zooming target centre using a given width and height margins (Figure 6-7).

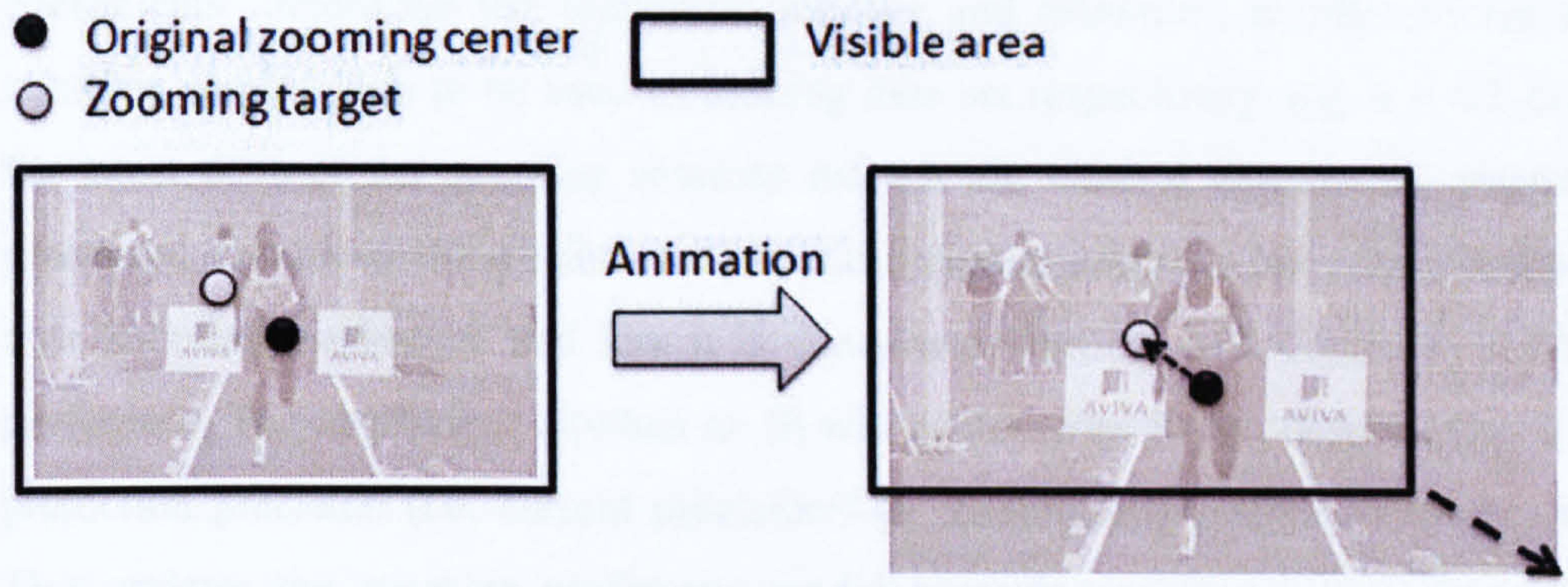


Figure 6-4 Zooming animation implementation

Time-shift playback and video quality adaptation is built upon the Internet Information Server (IIS) smooth streaming player development kit that provides APIs for using smooth streaming[7]. The implementation of personalised zooming, using the algorithm in Table 4-11 , consists of a set of control libraries executed in a sequential order using the personalised zooming control model proposed in Figure 4-18 and Figure 4-19. The class library names correspond to the key process names in the control models.

- *Personalisation* decision control handles the user input and distinguishes the personalization request and non-personalization request (see Figure 4-18). In the implementation, the personalization request or interaction trigger is defined as a single tap on the space bar, whereas the non-personalization request is defined as a double click on mouse left button.

[6] http://msdn.microsoft.com/en-us/library/bb509561%28VS.85%29.aspx
[7] http://www.iis.net/expand/SmoothStreaming

- *User Profile control* (see user profile processing process in Figure 4-19) is used to make either a personalization request) or non-personalization request.

- *Clustering control* implements the region of interest clustering process in Figure 4-19) based upon the algorithms in Table 4-11, Table 4-13 and Table 4-14. The predicted user zooming region central point coordinate is obtained from a Wilcoxon signed-rank (WSR) test. In order to mitigate the impact of significant user zooming preference changes caused by recent video content change versus when it is more stable across the use zooming sessions, three critical coefficients can be initially set. These are $0 < \alpha < 1$, $0 < \beta < 1$ ($\alpha > \beta$ and $\alpha- \beta > 0.05$) and $0 < \phi < 1$. The first two coefficients' determine the maximum number and minimum number of recorded zooming session data to be used as training data set respectively, e.g. $\alpha = 0.6$ means the latest 60% of the zooming sessions records are used. ϕ denotes the prediction precision. The lower the α value, the more the system adapts to the current session's user zooming preference and less it is concerned with a user's historical zooming preference. The coefficient α (when $\alpha > \beta$) will decrease by 5% by default when a low prediction precision (i.e. current precision< ϕ) continuously occurs for triple times. This updates the zooming preference model according to the changes in a user's recent zooming preference.

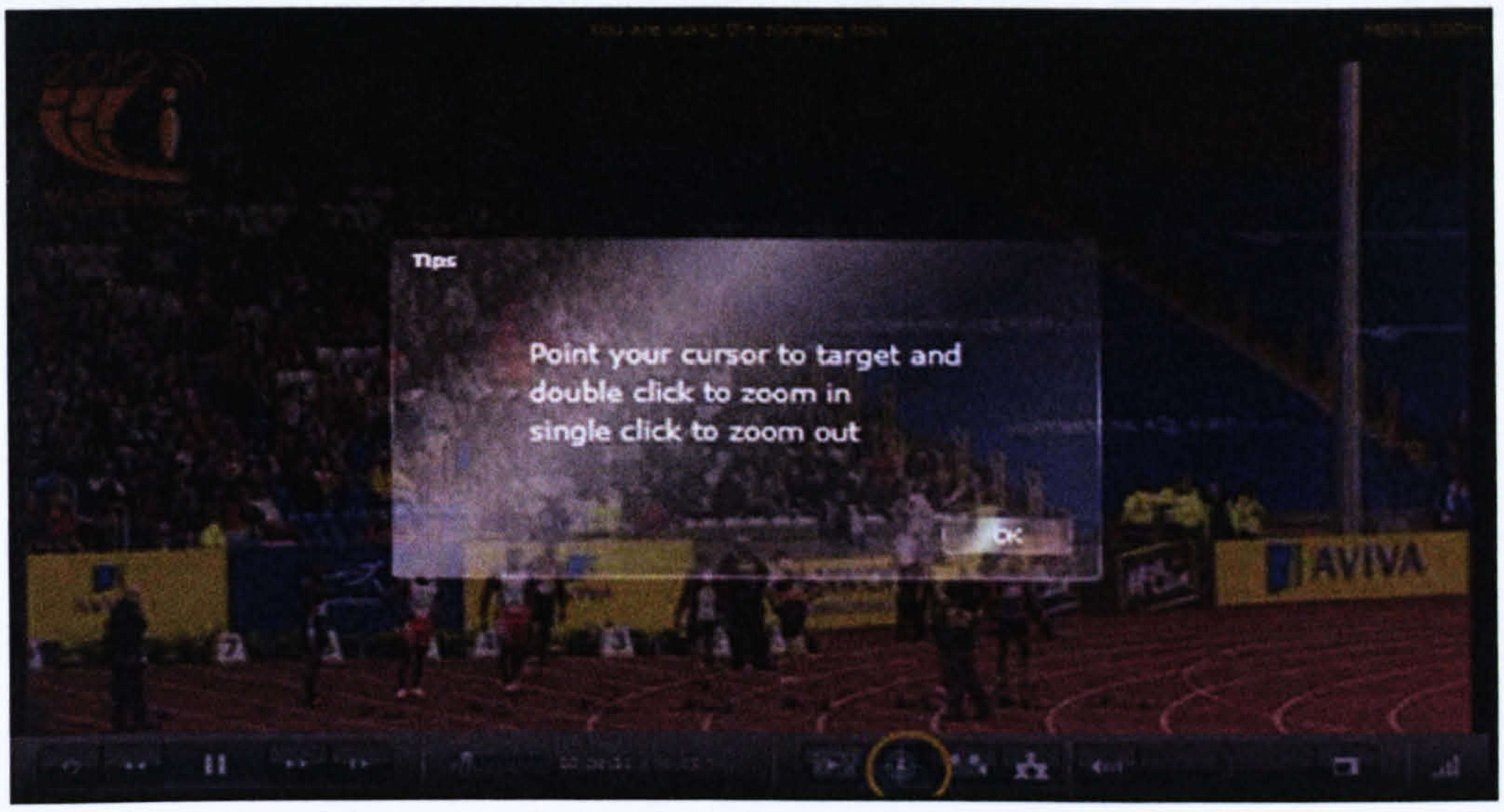


Figure 6-5 Video ZUI

The UI for selective target zooming (Figure 6-5) contains a pop up instruction window. The non-personalised zoom-in starts with a double click (clicks within 200 milliseconds) on mouse left button and zoom-out with a single click on mouse left button. The personalised zoom-in is triggered with a single press (press interval more than 500 milliseconds) on space bar and a second press triggers a zoom-out.

6.1.6 Time-Shift Viewing

The core components for time-shift viewing that are implemented are video analysis and content annotation (Figure 4-21.) The video analysis component implements the first 4 steps of the incidents highlighting model (Figure 4-24). A first .NET video processing class library implements the screen capture, colour channel conversion and motion data normality test. A screen capture approach is used to capture video screen images at a frequency determined by the frame rate of the video. The default capture frequency is set to be 40ms for a video fps of 25. The luminosity in the captured image is altered to a grey scale. A Kolmogorov–Smirnov test is used as the Goodness-Of-Fit test due to its robustness when dealing with small size sample data. The alpha value of this test is set to be 0.05 by default. Finally a z-score is obtained for the last motion value in the Gaussian distribution.

A second .NET control dealing with motion conditions determination (see Figure 4-23) generates the system actions including system editing decisions or system suggestions. The default critical z-score for left and right tail conditions are -1.64 and 1.65. They indicate the boundaries for 5% extreme data (i.e. smaller than -1.64 or greater than 1.65). In this implementation, the general rule for the motion condition (see Figure 4-23) and system actions rule is shown in Figure 6-6.

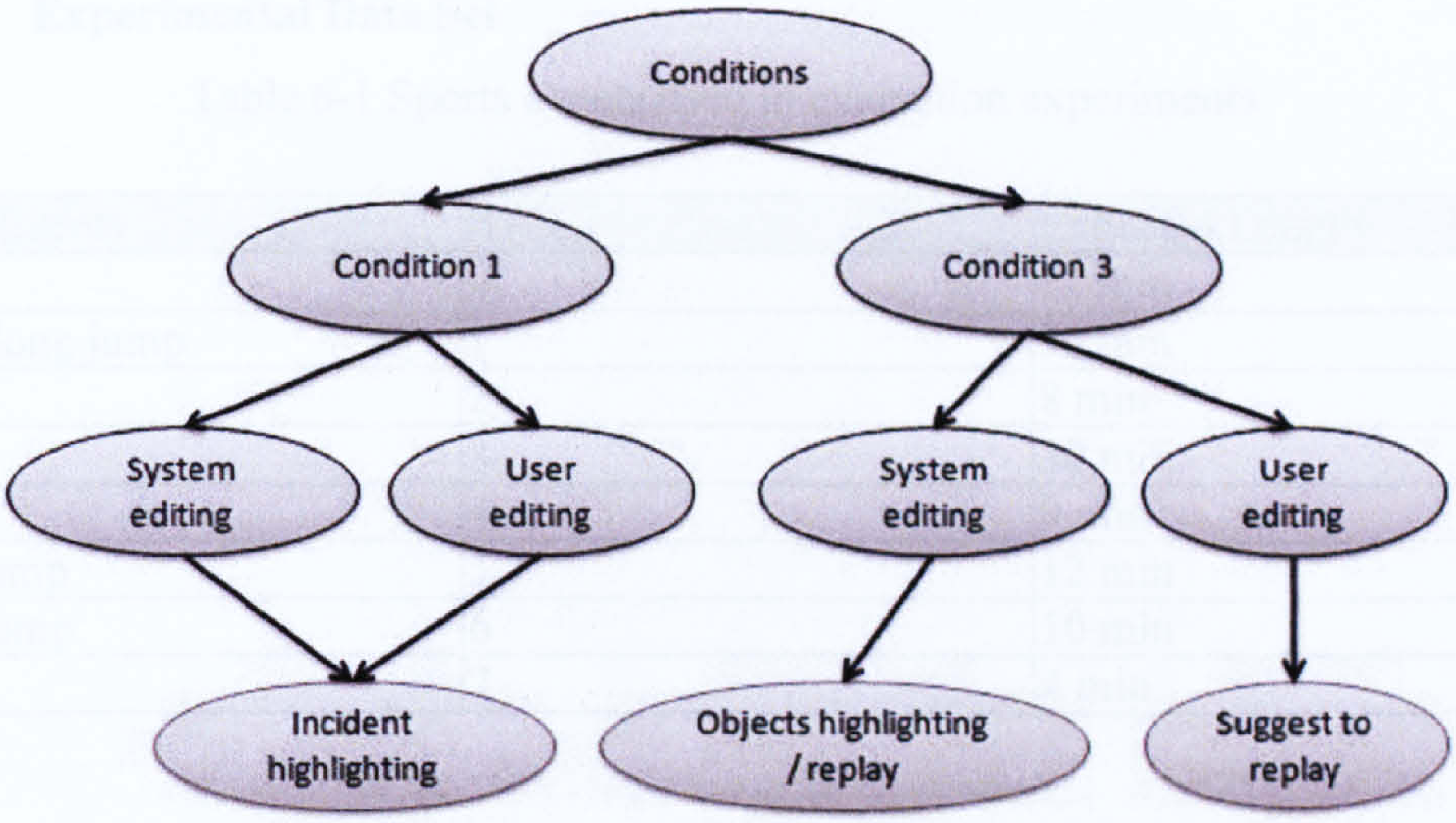


Figure 6-6 Decision rule of system actions according to the detected motion conditions

To annotate objects, a visual overlay is used to visually highlight objects given the boundary coordinates generated by a video analysis component. In this implementation, highlighting annotation is used for highlighting the fastest moving objects. The UI of visual annotation is shown in Figure 6-7.

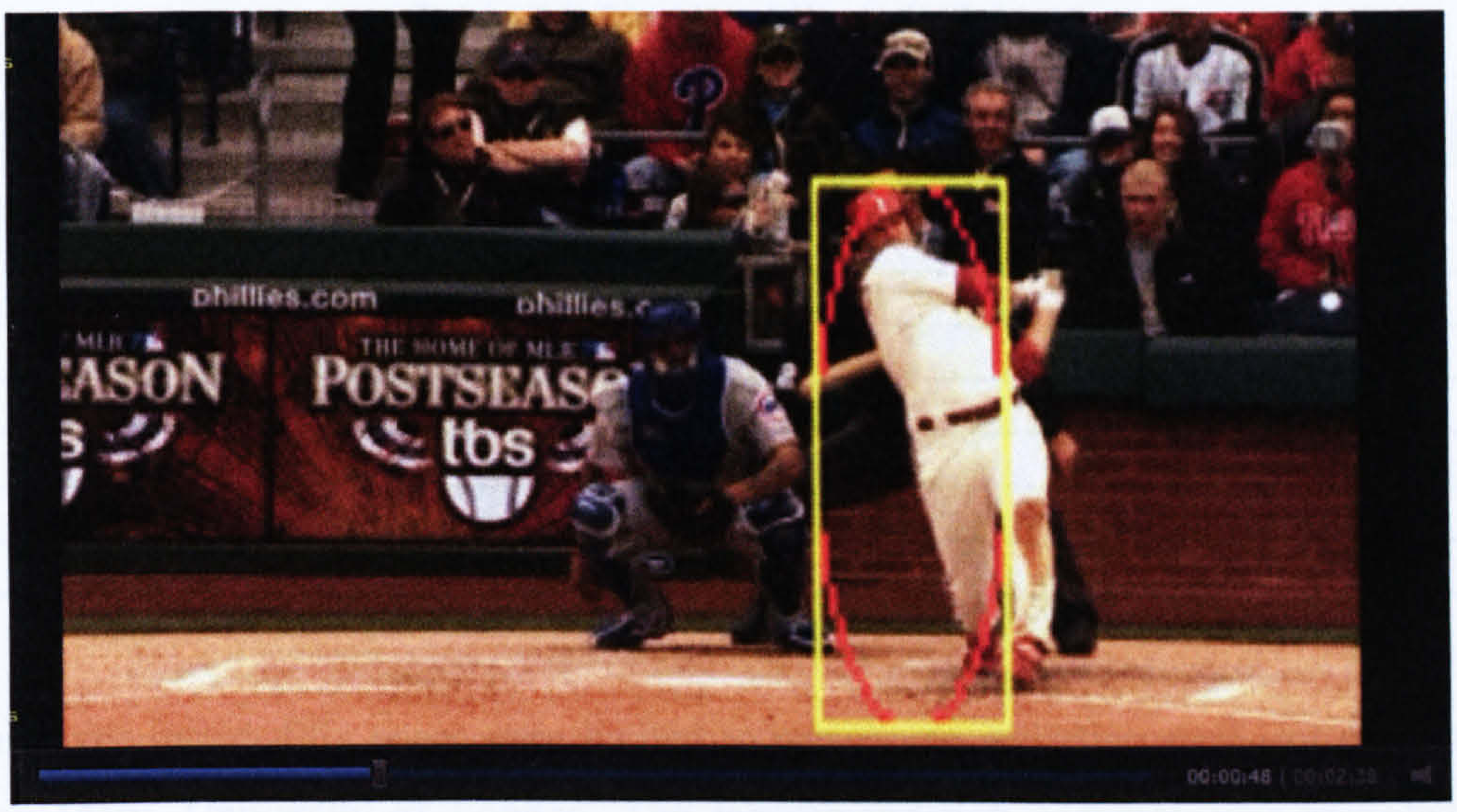


Figure 6-7 Object visual annotation UI

6.2 Personalisation Evaluation

6.2.1 Personalised Sports Events Selection (Individual Recommendations)

This section evaluates the personalised events browsing that recommending events to individual users.

6.2.1.1 Experimental Data Set

Table 6-1 Sports events used in evaluation experiments

Sports Events	ID	Expected Length
100m	0	12 min
Triple long jump	1	30 min
800m	2	8 min
5000m	3	30 min
Javelin	4	9 min
High jump	5	12 min
Long jump	6	10 min
400m	7	4 min

The test video content consists of eight different athletics sports events from Birmingham grand prix 2008 including 100m, long jump, triple long jump, javelin, high jump, 400m, 800m and 5000m. The live feed is simulated with a predefined schedule. The schedule advances each time a user opens the events selection menu. Table 6-1 shows the chosen events for the experiments with their IDs and expected length. A list of the first eight sessions of simulated live events schedule are shown as following.

Table 6-2 Simulated live schedule

Simulated Live Schedule Session	Event ID
1	0,4,5,7
2	1,2,3,4,7
3	0,3,5,6,7
4	3,4,5
5	5,6
6	2,3,4,5,7
7	0,1,2,3,4,6,7
8	0,2,4,6

In this experiment, fifty users were invited to participate including staff and students at Queen Mary, University of London. The recommendation accuracy was recorded for each user starting from the second event viewing session. At the end of the experiment, twelve users had viewed at least six events; the other users had viewed less than six events. The first five recommendation accuracy values from those twelve users who viewed more than six events were analysed in more detail. The RA values collected are listed in Table 6-3.

Table 6-3 12 Users' recommendation accuracy values

User ID	Recommendation Accuracy (personalised event list). Ordered by Schedule Session 2, 3, 4 , 5 and 6				
1	0.20	0.60	0.67	1.00	0.60
2	0.80	0.80	0.67	1.00	1.00
3	1.00	0.80	0.67	1.00	0.40
4	0.40	0.80	0.67	1.00	1.00
5	0.20	0.80	0.33	1.00	0.60
6	0.60	0.80	0.67	0.00	0.80
7	0.40	1.00	0.33	1.00	1.00
8	0.60	0.80	0.67	1.00	1.00
9	0.80	1.00	1.00	1.00	0.80
10	0.80	1.00	0.33	1.00	0.80
11	0.40	0.60	0.67	1.00	1.00
	Non Recommendation Accuracy (random event list) Ordered by Schedule Session 2, 3, 4 , 5 and 6				
1	0.20	0.40	0.33	1.00	1.00
2	0.60	0.60	0.33	0.00	0.20
3	0.40	0.60	0.67	0.00	0.20
4	0.40	0.80	0.67	1.00	0.20
5	0.40	0.60	0.33	0.00	0.20
6	0.20	0.40	0.33	1.00	0.60
7	0.20	1.00	0.33	0.00	1.00
8	0.20	0.20	0.33	1.00	0.40
9	0.60	0.20	0.33	1.00	0.40
10	1.00	0.20	0.33	0.00	0.20
11	0.60	0.20	0.33	1.00	0.80
12	0.60	0.20	0.33	1.00	0.20

6.2.1.2 Run-time Personalisation Evaluation

In order to evaluate the personalisation performance during personalisation, the recommendation accuracy median is used as a test statistic. In this evaluation, four dummy recommendation accuracy medians are added before the first recommendation accuracy in order to achieve the statistical accuracy[8]. Here, it is assumed that the 4 pseudo RA values have the same value as the RA in second viewing session. For example, if user's first recommendation accuracy is 0.20, the four pseudo RA will also be 0.20. Table 6-4 recompiles the Table 6-3 so that initial four pseudo RAs are added to each of the records.

[8] The minimum sample size for WSR tests in this experiment is >5

Table 6-4 Recommendation accuracy values for operational evaluation, shaded values are pseudo RA values for statistic calculation purpose

User ID	Recommendation Accuracy (personalised event list)									Non Recommendation Accuracy (random event list)								
1	0.20	0.20	0.20	0.20	0.20	0.60	0.67	1.00	0.60	0.20	0.20	0.20	0.20	0.20	0.40	0.33	1.00	1.00
2	0.80	0.80	0.80	0.80	0.80	0.80	0.67	1.00	1.00	0.60	0.60	0.60	0.60	0.60	0.60	0.33	0.00	0.20
3	1.00	1.00	1.00	1.00	1.00	0.80	0.67	1.00	0.40	0.40	0.40	0.40	0.40	0.40	0.60	0.67	0.00	0.20
4	0.40	0.40	0.40	0.40	0.40	0.80	0.67	1.00	1.00	0.40	0.40	0.40	0.40	0.40	0.80	0.67	1.00	0.20
5	0.20	0.20	0.20	0.20	0.20	0.80	0.33	1.00	0.60	0.40	0.40	0.40	0.40	0.40	0.60	0.33	0.00	0.20
6	0.60	0.60	0.60	0.60	0.60	0.80	0.67	0.00	0.80	0.20	0.20	0.20	0.20	0.20	0.40	0.33	1.00	0.60
7	0.40	0.40	0.40	0.40	0.40	1.00	0.33	1.00	1.00	0.20	0.20	0.20	0.20	0.20	1.00	0.33	0.00	1.00
8	0.60	0.60	0.60	0.60	0.60	0.80	0.67	1.00	1.00	0.20	0.20	0.20	0.20	0.20	0.20	0.33	1.00	0.40
9	0.80	0.80	0.80	0.80	0.80	1.00	1.00	1.00	0.80	0.60	0.60	0.60	0.60	0.60	0.20	0.33	1.00	0.40
10	0.80	0.80	0.80	0.80	0.80	1.00	0.33	1.00	0.80	1.00	1.00	1.00	1.00	1.00	0.20	0.33	0.00	0.20
11	0.40	0.40	0.40	0.40	0.40	0.60	0.67	1.00	1.00	0.60	0.60	0.60	0.60	0.60	0.20	0.33	1.00	0.80
12	1.00	1.00	1.00	1.00	1.00	0.80	0.67	1.00	0.80	0.60	0.60	0.60	0.60	0.60	0.20	0.33	1.00	0.20

A Wilcoxon signed-rank (WSR) test is used for the evaluation; the significance level i.e. p value is set to 0.05. The Hypothesis Test (H0) compares the recommendation accuracy produced from personalisation mechanism (see section 5.6.2)

In order to validate the performance of both approaches, the fixed value is set to be a fixed median value 0.6, 0.7 and 0.8, whereas the variable is set to be the up to date median of previous *n* recommendation accuracy values.

Table 6-5 Hypothesis testing with fixed RA median of 0.6 operational, p=0.05

User ID	Recommendation Accuracy (personalised event list) given the threshold median is 0.6								
1	0.20	0.20	0.20	0.20	0.20	**0.60** ($p<0.05$)	**0.67** ($p<0.05$)	1.00 ($p>0.05$)	0.60 ($p>0.05$)
2	0.80	0.80	0.80	0.80	0.80	0.80 ($p>0.05$)	0.67 ($p>0.05$)	1.00 ($p>0.05$)	1.00 ($p>0.05$)
3	1.00	1.00	1.00	1.00	1.00	0.80 ($p>0.05$)	0.67 ($p>0.05$)	1.00 ($p>0.05$)	0.40 ($p>0.05$)
4	0.40	0.40	0.40	0.40	0.40	0.80 ($p>0.05$)	0.67 ($p>0.05$)	1.00 ($p>0.05$)	1.00 ($p>0.05$)
5	0.20	0.20	0.20	0.20	0.20	**0.80** ($p<0.05$)	**0.33** ($p<0.05$)	1.00 ($p>0.05$)	0.60 ($p>0.05$)
6	0.60	0.60	0.60	0.60	0.60	0.80 ($p>0.05$)	0.67 ($p>0.05$)	0.00 ($p>0.05$)	0.80 ($p>0.05$)
7	0.40	0.40	0.40	0.40	0.40	1.00 ($p>0.05$)	0.33 ($p>0.05$)	1.00 ($p>0.05$)	1.00 ($p>0.05$)
8	0.60	0.60	0.60	0.60	0.60	0.80 ($p>0.05$)	0.67 ($p>0.05$)	1.00 ($p>0.05$)	1.00 ($p>0.05$)
9	0.80	0.80	0.80	0.80	0.80	1.00 ($p>0.05$)	1.00 ($p>0.05$)	1.00 ($p>0.05$)	0.80 ($p>0.05$)
10	0.80	0.80	0.80	0.80	0.80	1.00 ($p>0.05$)	0.33 ($p>0.05$)	1.00 ($p>0.05$)	0.80 ($p>0.05$)
11	0.40	0.40	0.40	0.40	0.40	**0.60** ($p<0.05$)	**0.67** ($p<0.05$)	1.00 ($p>0.05$)	1.00 ($p>0.05$)
12	1.00	1.00	1.00	1.00	1.00	0.80 ($p>0.05$)	0.67 ($p>0.05$)	1.00 ($p>0.05$)	0.80 ($p>0.05$)

Table 6-5 shows the test results using a fixed RA median value of 0.6, 4 tests out of 48 tests are rejected.

Table 6-6 Hypothesis testing with fixed RA median of 0.7 operational, p=0.05

User ID	Recommendation Accuracy (personalised event list) given the threshold median is 0.7								
1	0.20	0.20	0.20	0.20	0.20	**0.60** (p<0.05)	**0.67** (p<0.05)	**1.00** (p<0.05)	**0.60** (p<0.05)
2	0.80	0.80	0.80	0.80	0.80	0.80 (p>0.05)	0.67 (p>0.05)	1.00 (p>0.05)	1.00 (p>0.05)
3	1.00	1.00	1.00	1.00	1.00	0.80 (p>0.05)	0.67 (p>0.05)	1.00 (p>0.05)	0.40 (p>0.05)
4	0.40	0.40	0.40	0.40	0.40	**0.80** (p<0.05)	**0.67** (p<0.05)	1.00 (p>0.05)	1.00 (p>0.05)
5	0.20	0.20	0.20	0.20	0.20	**0.80** (p<0.05)	**0.33** (p<0.05)	**1.00** (p<0.05)	**0.60** (p<0.05)
6	0.60	0.60	0.60	0.60	0.60	0.80 (p>0.05)	0.67 (p>0.05)	**0.00** (p<0.05)	0.80 (p>0.05)
7	0.40	0.40	0.40	0.40	0.40	1.00 (p>0.05)	**0.33** (p<0.05)	1.00 (p>0.05)	1.00 (p>0.05)
8	0.60	0.60	0.60	0.60	0.60	0.80 (p>0.05)	0.67 (p>0.05)	1.00 (p>0.05)	1.00 (p>0.05)
9	0.80	0.80	0.80	0.80	0.80	1.00 (p>0.05)	1.00 (p>0.05)	1.00 (p>0.05)	0.80 (p>0.05)
10	0.80	0.80	0.80	0.80	0.80	1.00 (p>0.05)	0.33 (p>0.05)	1.00 (p>0.05)	0.80 (p>0.05)
11	0.40	0.40	0.40	0.40	0.40	**0.60** (p<0.05)	**0.67** (p<0.05)	1.00 (p>0.05)	1.00 (p>0.05)
12	1.00	1.00	1.00	1.00	1.00	0.80 (p>0.05)	0.67 (p>0.05)	1.00 (p>0.05)	0.80 (p>0.05)

Table 6-6 shows the testing results using a fixed RA median value of 0.7. 13 tests out of 48 tests are rejected.

Table 6-7 Hypothesis testing with fixed RA median of 0.8 operational, p=0.05

User ID	Recommendation Accuracy (personalised event list) given the threshold median is 0.8								
1	0.20	0.20	0.20	0.20	0.20	**0.60** ($p<0.05$)	**0.67** ($p<0.05$)	**1.00** ($p<0.05$)	**0.60** ($p<0.05$)
2	0.80	0.80	0.80	0.80	0.80	0.80 ($p>0.05$)	0.67 ($p>0.05$)	1.00 ($p>0.05$)	1.00 ($p>0.05$)
3	1.00	1.00	1.00	1.00	1.00	0.80 ($p>0.05$)	0.67 ($p>0.05$)	1.00 ($p>0.05$)	0.40 ($p>0.05$)
4	0.40	0.40	0.40	0.40	0.40	**0.80** ($p<0.05$)	**0.67** ($p<0.05$)	1.00 ($p>0.05$)	1.00 ($p>0.05$)
5	0.20	0.20	0.20	0.20	0.20	**0.80** ($p<0.05$)	**0.33** ($p<0.05$)	**1.00** ($p<0.05$)	**0.60** ($p<0.05$)
6	0.60	0.60	0.60	0.60	0.60	0.80 ($p>0.05$)	0.67 ($p>0.05$)	**0.00** ($p<0.05$)	**0.80** ($p<0.05$)
7	0.40	0.40	0.40	0.40	0.40	**1.00** ($p>0.05$)	**0.33** ($p<0.05$)	**1.00** ($p>0.05$)	**1.00** ($p>0.05$)
8	0.60	0.60	0.60	0.60	0.60	0.80 ($p>0.05$)	0.67 ($p>0.05$)	1.00 ($p>0.05$)	1.00 ($p>0.05$)
9	0.80	0.80	0.80	0.80	0.80	1.00 ($p>0.05$)	1.00 ($p>0.05$)	1.00 ($p>0.05$)	0.80 ($p>0.05$)
10	0.80	0.80	0.80	0.80	0.80	1.00 ($p>0.05$)	0.33 ($p>0.05$)	1.00 ($p>0.05$)	0.80 ($p>0.05$)
11	0.40	0.40	0.40	0.40	0.40	**0.60** ($p<0.05$)	**0.67** ($p<0.05$)	1.00 ($p>0.05$)	1.00 ($p>0.05$)
12	1.00	1.00	1.00	1.00	1.00	0.80 ($p>0.05$)	0.67 ($p>0.05$)	1.00 ($p>0.05$)	0.80 ($p>0.05$)

Table 6-7 shows the testing results using a fixed RA median value of 0.8. 18 tests out of 48 tests are rejected.

To conclude this suggests that the use of a higher fixed RA threshold value can lead to higher hypothesis test rejections.

Table 6-8 Hypothesis testing with previous RA median, p=0.05

User ID	Recommendation Accuracy (personalised event list)								
1	0.20	0.20	0.20	0.20	0.20	0.60 (p>0.05)	0.67 (p>0.05)	1.00 (p>0.05)	0.60 (p>0.05)
2	0.80	0.80	0.80	0.80	0.80	0.80 (p>0.05)	**0.67** (p<0.05)	1.00 (p>0.05)	1.00 (p>0.05)
3	1.00	1.00	1.00	1.00	1.00	**0.80** (p<0.05)	**0.67** (p<0.05)	1.00 (p>0.05)	**0.40** (p<0.05)
4	0.40	0.40	0.40	0.40	0.40	0.80 (p>0.05)	0.67 (p>0.05)	1.00 (p>0.05)	1.00 (p>0.05)
5	0.20	0.20	0.20	0.20	0.20	0.80 (p>0.05)	0.33 (p>0.05)	1.00 (p>0.05)	0.60 (p>0.05)
6	0.60	0.60	0.60	0.60	0.60	0.80 (p>0.05)	0.67 (p>0.05)	**0.00** (p<0.05)	0.80 (p>0.05)
7	0.40	0.40	0.40	0.40	0.40	1.00 (p>0.05)	**0.33** (p<0.05)	1.00 (p>0.05)	1.00 (p>0.05)
8	0.60	0.60	0.60	0.60	0.60	0.80 (p>0.05)	0.67 (p>0.05)	1.00 (p>0.05)	1.00 (p>0.05)
9	0.80	0.80	0.80	0.80	0.80	1.00 (p>0.05)	1.00 (p>0.05)	1.00 (p>0.05)	0.80 (p>0.05)
10	0.80	0.80	0.80	0.80	0.80	1.00 (p>0.05)	**0.33** (p<0.05)	1.00 (p>0.05)	0.80 (p>0.05)
11	0.40	0.40	0.40	0.40	0.40	0.60 (p>0.05)	0.67 (p>0.05)	1.00 (p>0.05)	1.00 (p>0.05)
12	1.00	1.00	1.00	1.00	1.00	**0.80** (p<0.05)	**0.67** (p<0.05)	1.00 (p>0.05)	**0.80** (p<0.05)

Table 6-8 presents the result of Hypothesis Testing for each user during personalisation. The bold numbers indicate that the rejection of hypothesis. It is noted that the higher the fixed RA median is, the higher the hypothesis rejection rate is for some users. For example, user 1 and user 5 have a 100% rejection rate when the threshold RA median is set to 0.7 and 0.8, but when the threshold RA median is set to 0.6 both had a rejection rate of 50%. The average hypothesis rejection rates of involved users for the fixed RA median are 50%[9], 58.3% and 75%. When the RA median is set to be a variable, i.e. based upon the previous RA median values, the average hypothesis rejection rate for involved user is 41.7%.

The implication of this is that the use of Hypothesis Test at run time can be challenging when the recommendation accuracy threshold value is a fixed value as it is difficult to determine such value. It has a high hypothesis rejection rate when such value is set to be

[9] This calculation is based upon users: 1, 5 &11, each has two out four hypothesis rejection records, (2/4+2/4+2/4)/3 =50%

relatively high. In addition, a 'crisp values' can easily hide the fact that the recommendation accuracy for some users in effect increase over time, e.g., user 11. An alternative approach is to use a recommendation accuracy threshold value based upon previous recommendation accuracy values. This tends to generate a smaller hypothesis rejection rate and is able to allow the system to adaptively do the Hypothesis Testing, e.g. in the case of user 11, the hypothesis is always accepted when the recommendation accuracy gradually increases. This can be used to allow the system to adjust itself to perform the personalisation when browsing subsequent events. For example:

Rule 6.2.1.2-1: If the hypothesis is rejected, the system can update the current recommendation accuracy by reducing the current recommendation accuracy variable by 15% given the recommendation accuracy is greater than 15%, otherwise the value will be 0.

Table 6-9: Apply rule 6.2.1.2-1 to user 12 to decide to update the current recommendation accuracy threshold value

User 12 without rule	1.00	1.00	1.00	1.00	1.00	**0.80** ($p<0.05$)	**0.67** ($p<0.05$)	1.00 ($p>0.05$)	**0.80** ($p<0.05$)
User 12 (with rule)	1.00	1.00	1.00	1.00	1.00	**0.80 To 0.65** ($p<0.05$)	**0.65 To 0.51** ($p<0.05$)	1.00 ($p>0.05$)	0.80 ($p>0.05$)

With respect to user 12 (Table 6-3), the hypothesis rejection rate will be reduced from 75% to 66% as shown in Table 6-9. **This result therefore suggests that a rule based RA threshold is able to reduce the hypothesis test rejection rate.**

Hypothesis Testing can be used to compare the results produced from a system that uses personalisation against the results produced by a system without non-personalisation based upon H1 (see section 5.6.2).

Table 6-10 Hypothesis testing with previous recommendation accuracy produced by random approach, p=0.05

User ID	Recommendation Accuracy (personalised event list)								
1	0.20	0.20	0.20	0.20	0.20	0.60 (p>0.05)	0.67 (p>0.05)	1.00 (p>0.05)	0.60 (p>0.05)
2	0.80	0.80	0.80	0.80	0.80	0.80 (p>0.05)	0.67 (p>0.05)	1.00 (p>0.05)	1.00 (p>0.05)
3	1.00	1.00	1.00	1.00	1.00	0.80 (p>0.05)	0.67 (p>0.05)	1.00 (p>0.05)	0.40 (p>0.05)
4	0.40	0.40	0.40	0.40	0.40	0.80 (p>0.05)	0.67 (p>0.05)	1.00 (p>0.05)	1.00 (p>0.05)
5	0.20	0.20	0.20	0.20	0.20	0.80 (p>0.05)	**0.33** (p<0.05)	1.00 (p>0.05)	0.60 (p>0.05)
6	0.60	0.60	0.60	0.60	0.60	0.80 (p>0.05)	0.67 (p>0.05)	**0.00** (p<0.05)	0.80 ((p>0.05))
7	0.40	0.40	0.40	0.40	0.40	1.00 (p>0.05)	0.33 (p>0.05)	1.00 (p>0.05)	1.00 (p>0.05)
8	0.60	0.60	0.60	0.60	0.60	0.80 (p>0.05)	0.67 (p>0.05)	1.00 (p>0.05)	1.00 (p>0.05)
9	0.80	0.80	0.80	0.80	0.80	1.00 (p>0.05)	1.00 (p>0.05)	1.00 (p>0.05)	0.80 (p>0.05)
10	0.80	0.80	0.80	0.80	0.80	1.00 (p>0.05)	**0.33** (p<0.05)	1.00 (p>0.05)	0.80 (p>0.05)
11	0.40	0.40	0.40	0.40	0.40	0.60 (p>0.05)	0.67 (p>0.05)	1.00 (p>0.05)	1.00 (p>0.05)
12	1.00	1.00	1.00	1.00	1.00	0.80 (p>0.05)	0.67 (p>0.05)	1.00 (p>0.05)	0.80 (p>0.05)

With reference to the data in Table 6-4, Table 6-10 shows the test results. The hypothesis rejection occurs for three users (25% of the users). A rejection rate of 25% is obtained. **This suggests that the use of the personalisation mechanism is effective during the system execution compared to the situation without the use of personalisation.**

The results of this Hypothesis Testing can also be used to allow the system to adjust itself to perform personalisation in subsequent events browsing session. For example:

Rule 6.2.1.2-2: If the hypothesis is rejected continuously, e.g. 3 sequential times, the system can hide the personalisation from user in the next events browsing session, i.e. user will not see the personalised event list. The personalisation activates only when the hypothesis is not rejected in the next events browsing session

In this experiment, no continuous hypothesis rejection has occurred. In order to demonstrate the use of this rule, the 'continuously hypothesis rejection' condition is ignored. Hence, in the case of user 5, the system self-adjustment is the following:

Table 6-11: Apply Rule 6.2.1.2-2 to user 5 to decide to activate the personalisation

User 5 (without rule)	0.20	0.20	0.20	0.20	0.20	0.80 ($p>0.05$)	**0.33** ($p<0.05$)	1.00 ($p>0.05$)	0.60 ($p>0.05$)
User 5 (with rule)	0.20	0.20	0.20	0.20	0.20	0.80 ($p>0.05$)	**0.33** ($p<0.05$)	1.00 ($p>0.05$)	0.60 ($p>0.05$)

In the table above, when the RA reaches 0.33 ($p<0.05$), the personalisation will be inactivated in the next session, when the RA reaches 1.00 ($p>0.05$), the personalisation will be reactivated. **As a result, it can be concluded that adjustment of the personalisation can be achieved by using a rule-based approach.**

6.2.1.3 Post-Operation Testing Results

Two different Hypotheses are tested here are H2 and H4. H2 is to test personalisation consistency and H4 is to test personalisation scalability.

For the consistency test, a normality test is conducted using the Jarque-Bera test. Data with a normal distribution are used for an F-test, both with a p value of 0.05. For the non-normally distributed sample data, a WSR test will be used to test whether the median of these data is greater or equal than the average median of the consistent sample data.

Based upon the data in Table 6-3, ten users' data (user 1, 2, 3, 4, 5, 6, 8, 10, 11 and12) are normally distributed. The consistency rate based upon an F-test for these users is 100% with an average recommendation accuracy median of 73.3%. In addition, The prediction precision (PP) median value WSR test (left-tailed test)[10] of two groups of non-normally distributed user data (user 7 and 9) also indicate that the hypothesis is not rejected when the given median value of 73.3% with p value of 0.05. **Hence, based upon hypothesis testing, it can be concluded that the consistency rate of personalisation across these twelve users is 83.33% which suggests that personation is consistent.**

[10] The test is based upon the null hypothesis that the PP median of normally distributed data groups is less or equal to the testing user's median value

For the scalability test, a Spearman's rank correlation coefficient test is used. Table 6-12 shows the results of the correlation of the use with recommendation accuracy for each user.

Table 6-12 Correlation coefficient for number of uses and recommendation accuracy

User ID	Correlation Coefficient ρ	Correlation Coefficient Significance (left tail, p=0.05)
1	0.56	1.00 (>0.05)
2	0.63	1.00 (>0.05)
3	**-0.56**	1.00 (>0.05)
4	0.87	1.00 (>0.05)
5	0.50	1.00 (>0.05)
6	0.15	1.00 (>0.05)
7	0.45	1.00 (>0.05)
8	0.87	1.00 (>0.05)
9	0.00	1.00 (>0.05)
10	0.21	1.00 (>0.05)
11	0.97	1.00 (>0.05)
12	**-0.21**	1.00 (>0.05)

The results clearly indicate that ten of twelve users have positive correlation coefficient values. If defines the correlations as: [0.5 ~ 1]: strong positive correlation, [-1 ~ -0.5] strong negative correlation), (0 ~ 0.5) weak positive correlation, and (-0.5 ~ 0) weak negative correlation. Half of the users have a strong positive correlation coefficient (i.e. 50%). One user has a strong negative correlation coefficient (i.e. 1/12 = 8.33%). One user has a weak negative correlation coefficient (i.e. 8.33%) and four users have a weak positive correlation coefficient (i.e. 33.33%). **Therefore, it is concluded that the hypothesis is rejected and that personalisation is scalable.**

6.2.2 Personalised Sports Event Selection (Group Recommendations)

This section evaluates personalised events browsing where events are recommended to groups of users.

6.2.2.1 Experimental Data Set

The video content used and the users selected here is the same way as in section 6.2.1.1. Basic demographic information which includes user's gender, age, nationality and watching sports TV behaviour were obtained from pre-trial questionnaire. All collected user data was used as the ground truth for validating the proposed personalization. In this experiment, fifteen users were selected for evaluation as they had viewed at least four events. Therefore, their first three recommendation accuracy values can be studied given recommendation starts from second viewing session.

6.2.2.2 Run-time Personalisation Evaluation

In this evaluation, the Hypothesis Testing was used to assess hypothesis H1. The process of conducting operational evaluation is described in the following four steps. First, randomly select three users out of fifteen users as the test users for each viewing session. Second, use the other twelve users as the training set and generate the three personalised event lists for the selected test group of users. Third, calculate the average recommendation accuracy for the test users based upon any selections they make. Fourth, repeat the first three steps ten times.

Table 6-13, Table 6-14 and Table 6-15 show the recommendation accuracy results of recommendation generation using four methods (DT, BN, BMP and Random) across ten testing sessions for a schedule with session 2, 3 and 4 respectively.

Table 6-13 Recommendation accuracy results for schedule session 2

Testing Session (Schedule Session 2)	1	2	3	4	5	6	7	8	9	10
DT (%)	80	80	60	60	80	60	80	60	100	80
BN (%)	60	50	70	30	50	60	50	20	60	70
BPM (%)	60	50	60	50	60	70	80	60	70	70
Random (%)	50	20	60	20	40	40	60	20	60	20

The results from Table 6-13 indicate that the **group recommendation system has a higher recommendation accuracy (63% on average) than a system that does not use group recommendations (39% on average).** Among the three recommendation generation methods, the decision tree method has the highest average recommendation accuracy (mean = 74%, median = 80%). BPM has a higher average recommendation accuracy (mean = 63%, median = 60%) than BN (mean = 52%, median = 55%).

Table 6-14 Recommendation accuracy results for schedule session 3

Testing Session (Schedule Session 3)	1	2	3	4	5	6	7	8	9	10
DT (%)	70	60	80	80	80	80	100	100	80	80
BN (%)	50	60	70	60	70	50	50	60	60	70
BPM (%)	60	70	60	60	60	70	70	70	70	80
Random (%)	50	20	50	50	50	60	20	50	20	60

The results from Table 6-14 also indicate that **a group recommendation system has a higher recommendation accuracy (69.3% by average) than a system that does not use recommendations (43% by average).** Among the three recommendation generation methods, the decision tree has the highest average recommendation accuracy (mean = 81%, median = 80%). BPM has a higher average recommendation accuracy (mean = 67%, median = 70%) than BN (mean = 60%, median = 60%).

Table 6-15 Recommendation accuracy results for schedule session 4

Testing Session (Schedule Session 4)	1	2	3	4	5	6	7	8	9	10
DT (%)	66	33	100	67	67	67	67	100	33	100
BN (%)	33	33	33	33	67	33	33	33	67	67
BPM (%)	67	100	67	67	67	100	67	67	33	33
Random (%)	33	33	33	33	33	33	67	67	67	67

The results from Table 6-15 indicate that **some group recommendation system has a lower recommendation accuracy than the system that does not use recommendations.** Among the three recommendation generation methods, the decision tree still has the highest average recommendation accuracy (mean = 70%, median = 67%). BPM has a higher average recommendation accuracy (mean = 67%, median = 70%) than the system without personalisation. BN has the lowest average recommendation accuracy (mean = 43%, median = 33%).

Using the data from Table 6-13 Table 6-14 and Table 6-15, Hypothesis Testing can be started for the fifth[11] resample with WSR test, p=0.05.

Table 6-16 Hypothesis testing for a single personalization session with 10 X resampling (Schedule session 2)

Testing Session (Schedule Session 2)	1	2	3	4	5	6	7	8	9	10
DT (%)	80	80	60	60	80	60 (p>0.05)	*80* (p>0.05)	60 (p>0.05)	*100* (p>0.05)	*80* (p>0.05)
BN (%)	60	50	70	30	50	60 (p>0.05)	50 (p>0.05)	**20** (p<0.05)	60 (p>0.05)	70 (p>0.05)
BPM (%)	60	50	60	50	60	*70* (p>0.05)	*80* (p>0.05)	60 (p>0.05)	70 (p>0.05)	70 (p>0.05)
Random (%)	50	20	60	20	40	40 (p>0.05)	60 (p>0.05)	**20** (p<0.05)	60 (p>0.05)	**20** (p<0.05)

Table 6-17 Hypothesis Testing for a Single Personalization Session with 10 X resampling (Schedule session 3)

Testing Session (Schedule Session 3)	1	2	3	4	5	6	7	8	9	10
DT (%)	70	60	80	80	80	*80* (p>0.05)	*100* (p>0.05)	*100* (p>0.05)	*80* (p>0.05)	*80* (p>0.05)
BN (%)	50	60	70	60	70	50 (p>0.05)	50 (p>0.05)	60 (p>0.05)	60 (p>0.05)	70 (p>0.05)
BPM (%)	60	70	60	60	60	70 (p>0.05)	70 (p>0.05)	70 (p>0.05)	70 (p>0.05)	*80* (p>0.05)
Random (%)	50	20	50	50	50	60 (p>0.05)	**20** (p<0.05)	50 (p>0.05)	**20** (p<0.05)	60 (p>0.05)

[11] The minimum sample size for WSR tests in this experiment is >5

Table 6-18 Hypothesis testing for a single personalization session with 10 X resampling (Schedule session 4)

Testing Session (Schedule Session 4)	1	2	3	4	5	6	7	8	9	10
DT (%)	67	33	100	67	67	67 ($p>0.05$)	*67* ($p>0.05$)	*100* ($p>0.05$)	**33** ($p<0.05$)	*100* ($p>0.05$)
BN (%)	33	33	33	33	67	33 ($p>0.05$)	33 ($p>0.05$)	33 ($p>0.05$)	*67* ($p>0.05$)	*67* ($p>0.05$)
BPM (%)	67	100	67	67	67	*100* ($p>0.05$)	*67* ($p>0.05$)	67 ($p>0.05$)	**33** ($p<0.05$)	**33** ($p<0.05$)
Random (%)	33	33	33	33	33	33 ($p>0.05$)	*67* ($p>0.05$)	67 ($p>0.05$)	*67* ($p>0.05$)	*67* ($p>0.05$)

In this experiment, the group recommender system is tested whether it is able to choose the best recommendation generation method among provided with the rule below:

Rule 6.2.1.2-1: If the hypothesis is rejected, the recommendation generation method will be marked as -1. If the hypothesis is not rejected, the method with the largest recommendation accuracy is marked as +1; otherwise it is marked as 0. The candidate recommendation generation method can be obtained through ranking the sum of the score of each tested method at the end of Hypothesis Testing.

By applying the Rule 6.2.1.2-1 to Table 6-16, Table 6-17 and Table 6-18, it is clear that for schedule session 2, DT has the highest score which is +3. For the scheduled session 3, DT still has the highest score of 5. For the scheduled session 4, the system without personalization support has the highest score of +3. As a result, DT should be used for both schedule session 3 and schedule session 4. This result is in line with the ground truth as the DT generated the highest average recommendation accuracy (Table 6-17). **As a result, it is clear that the group recommender system is able to choose the best recommendation generation method using rules.**

6.2.2.3 Post-Operation Testing Results

The evaluation here is mainly to assess hypothesis H4. Based upon the collected data from fifty users on the second viewing session (i.e. schedule session 2), the number of the users is increased to 10, 15, 20, 25, 30, 35, 40, 45 and 50 by design. For each increment, 90% of the user population will be the training set and the remaining 10% is used as the test set. With the obtained recommendation accuracy of nine tests as shown in Table 6-19, a spearman's rank correlation coefficient test can be further conducted.

Table 6-19 Mean recommendation accuracy with corresponding user number

Total User No.	Mean Accuracy(ε)% (DT)	Mean Accuracy(ε)% (BN)	Mean Accuracy(ε)% (BMP)
10	100	20	60
15	80	50	70
20	70	50	60
25	60	40	53.3
30	76.7	46.7	60
35	80	65	70
40	80	70	75
45	84	64	72
50	88	60	74

The correlation coefficient tests give coefficient results of 0.20, 0.71 and 0.70 for DT, BN and BMP respectively. The corresponding left-tail significance tests (p=0.01) further confirm that all of the obtained coefficients are greater or equal to zero. **Therefore, it can be concluded that the hypothesis is not rejected and that personalisation is scalable.**

6.2.3 Personalised Multi-Angle Viewing

Here the multi-stream adaptation capability and the quality of personalisation were evaluated.

6.2.3.1 Experimental Data Set

The multi-stream adaptation performance was tested in a laboratory setting. An open source testing video stream was used for testing. Videos were encoded in the IIS smooth streaming format at different visual quality levels. Video streams were looped to simulate a live mode. Camera switching personalisation was evaluated with users. The test video content used for personalisation evaluation is the men's 5000m event in Aviva European Trials and UK Championships 2008 video stream (total length of 16 minutes). The supporting camera settings are shown in Figure 6-8.

For the personalisation evaluation, a combination of six expert and novice users participated in the test. At the beginning, they were shown screenshots corresponding to different camera views before choosing the camera switching sequence in the prototype system as shown in Figure 6-8. Users were asked to perform more than five camera switches (in each use session and to re-use the system at least six times) so that all cameras can be switched around. The prediction precision (PP) was directly generated based upon user interaction. In the meantime, the system also generates a random prediction on camera switching intervals representing the result without personalisation. These retrieved data are shown in Table 6-20.

Table 6-20 Retrieved PP for a camera switching system with personalisation versus a generated random PP on system without personalisation

User ID	Prediction Precision (with personalisation). Ordered by Task Use Session 2, 3, 4, 5 and 6					Random Prediction Precision (without personalisation) Ordered by Task Use Session 2, 3, 4, 5 and 6				
1	0.75	0.25	0.50	0.50	0.75	0.25	0.00	0.00	0.25	0.25
2	0.25	0.50	0.50	0.50	0.75	0.00	0.25	0.25	0.25	0.25
3	0.75	0.75	0.75	0.75	0.50	0.25	0.00	0.25	0.00	0.00
4	1.00	0.75	0.50	0.50	0.75	0.25	0.25	0.00	0.50	0.00
5	0.75	0.25	0.25	0.25	0.25	0.00	0.25	0.50	0.00	0.25
6	0.25	0.25	0.75	0.50	0.50	0.25	0.00	0.25	0.25	0.00

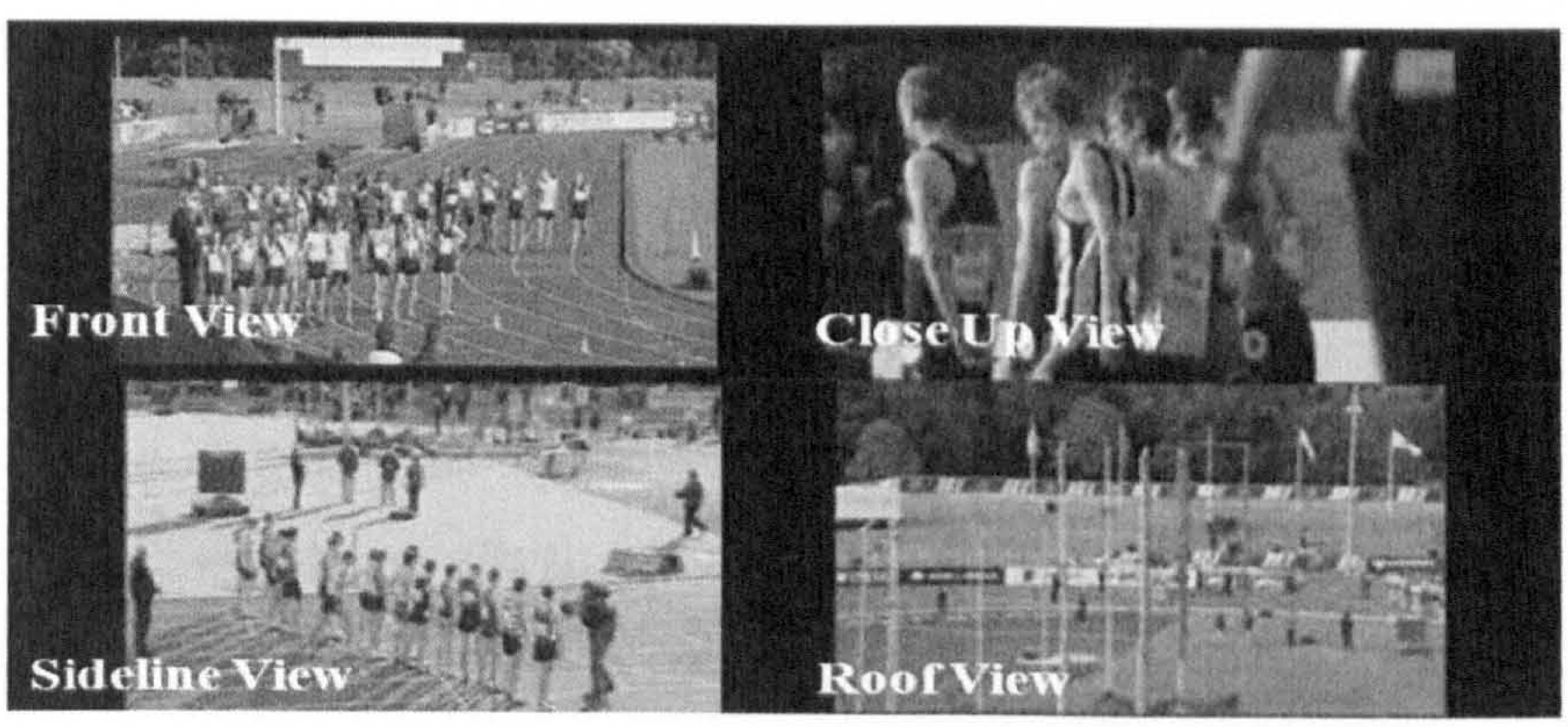


Figure 6-8 Supporting camera settings

6.2.3.2 Personalisation Task Interface Evaluation

In this experiment, the multi-stream adaptation performance is evaluated in terms of its capability to mitigate the bandwidth competition problem. A comparison approach is taken, i.e. a system with multi-stream adaptation versus a system without multi-stream adaptation. The video stream download speed is capped[12] to 2.5Mbps for the testing Web browser.

[12] This is achieved by using the NetLimiter at http://www.netlimiter.com/index.php

A sample video clip was encoded with the following bitrates: 300Kbps, 427Kbps, 608Kbps, 866Kbps, 1.23Mbps, 1.64Mbps and 2.43Mbps. IIS smooth stream technology was used to stream the video. Three copies of the same streams rendered with same video height and width are sequentially streamed by the system. The second and third streams were added 35 and 40 seconds respectively after the first stream was played. The first stream was used as the master stream in multi-stream adaptation case. Both tests were conducted under the same network connection conditions. The master stream's threshold value was set to be 0.5 (i.e. critical playback bitrates is 608Kbps). The playback bitrates of three streams were recorded at second intervals during the test. The testing period for both tests was 80 seconds. The obtained experimental results are illustrated in Figure 6-9 and Figure 6-10. The shaded regions for the system without multi-stream adaptation indicate the bandwidth competition phases, whereas the shaded area for the system with multi-stream adaptation indicates the adaptation phases.

With respect to Figure 6-9 and Figure 6-10, after 32s, the system with multi-stream adaptation started to adapt the master stream through reducing the playback bitrate to 1.64Kbps according to its video screen resolution whereas the system without multi-stream adaptation kept the maximum playback bitrate for the video stream at 2.43Mbp.

After 35s, a second stream was added to the system. A bandwidth competition problem immediately occurred in the system without multi-stream adaptation. The master stream playback bitrates plummeted and is companied with a sharp increase in the second stream playback bitrates. In the system with multi-stream adaptation, the system decreased the master stream playback bitrate automatically to a critical playback bitrate and increased the incoming stream playback bitrates to 1.64Mbps according to the bitrates allocation algorithm (i.e. case 2 in Table 4-8).

After 40s, a third stream was added to the system. The system with multi-stream adaptation immediately adjusts the second stream's playback bitrates through stepping down one visual quality level to 1.23Mbps. This allows the third incoming stream to use the lowest playback bitrates 300Kbps (i.e. case 4 in Table 4-8). The system without multi-stream adaptation is very sensitive to the incoming stream and the bandwidth competition problem occurs frequently afterwards. The master stream playback bitrate drops drastically again and third stream playback bitrates also decreased shortly afterwards.

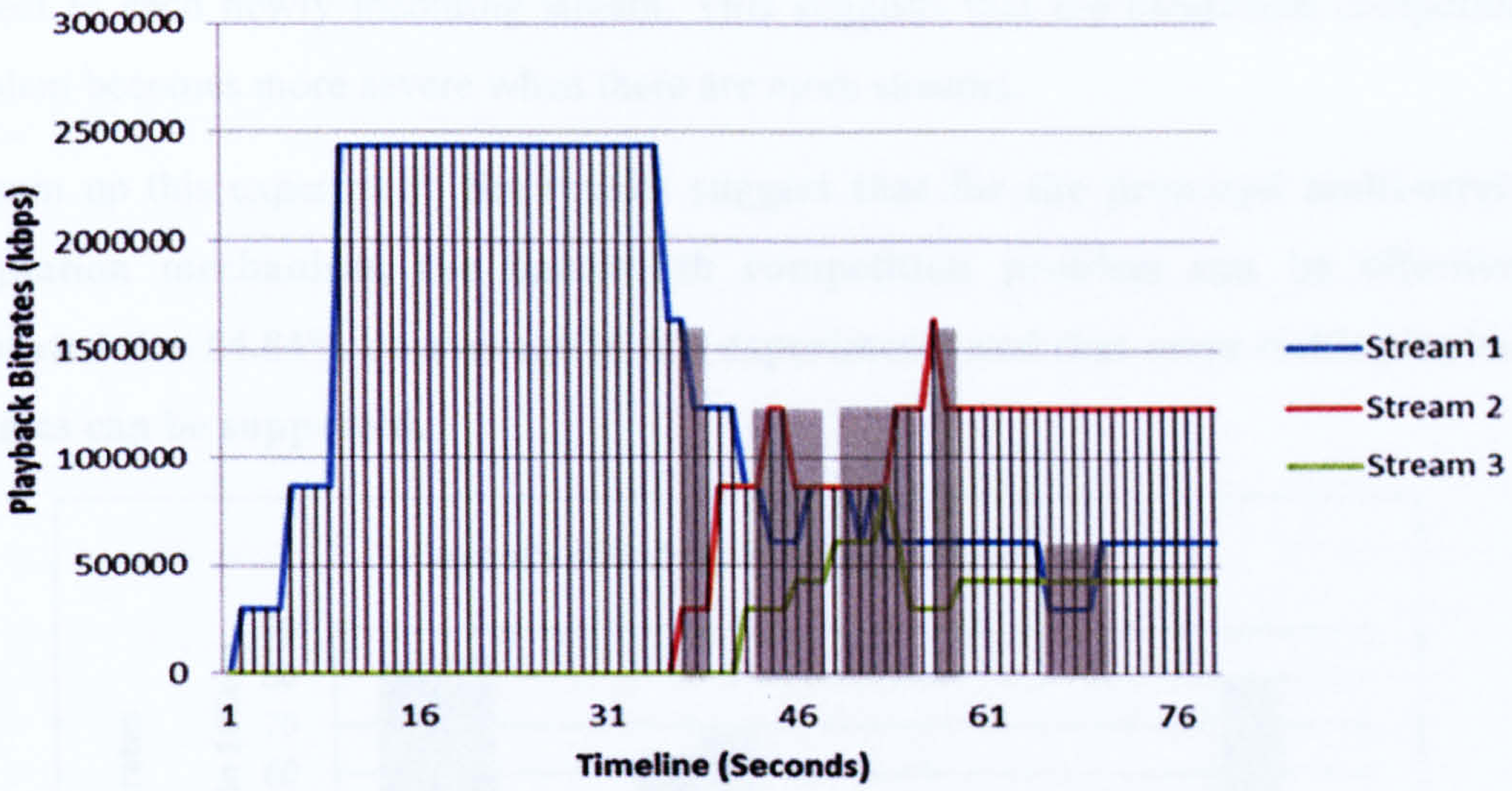


Figure 6-9 Bitrates changes without multi-stream adaptation

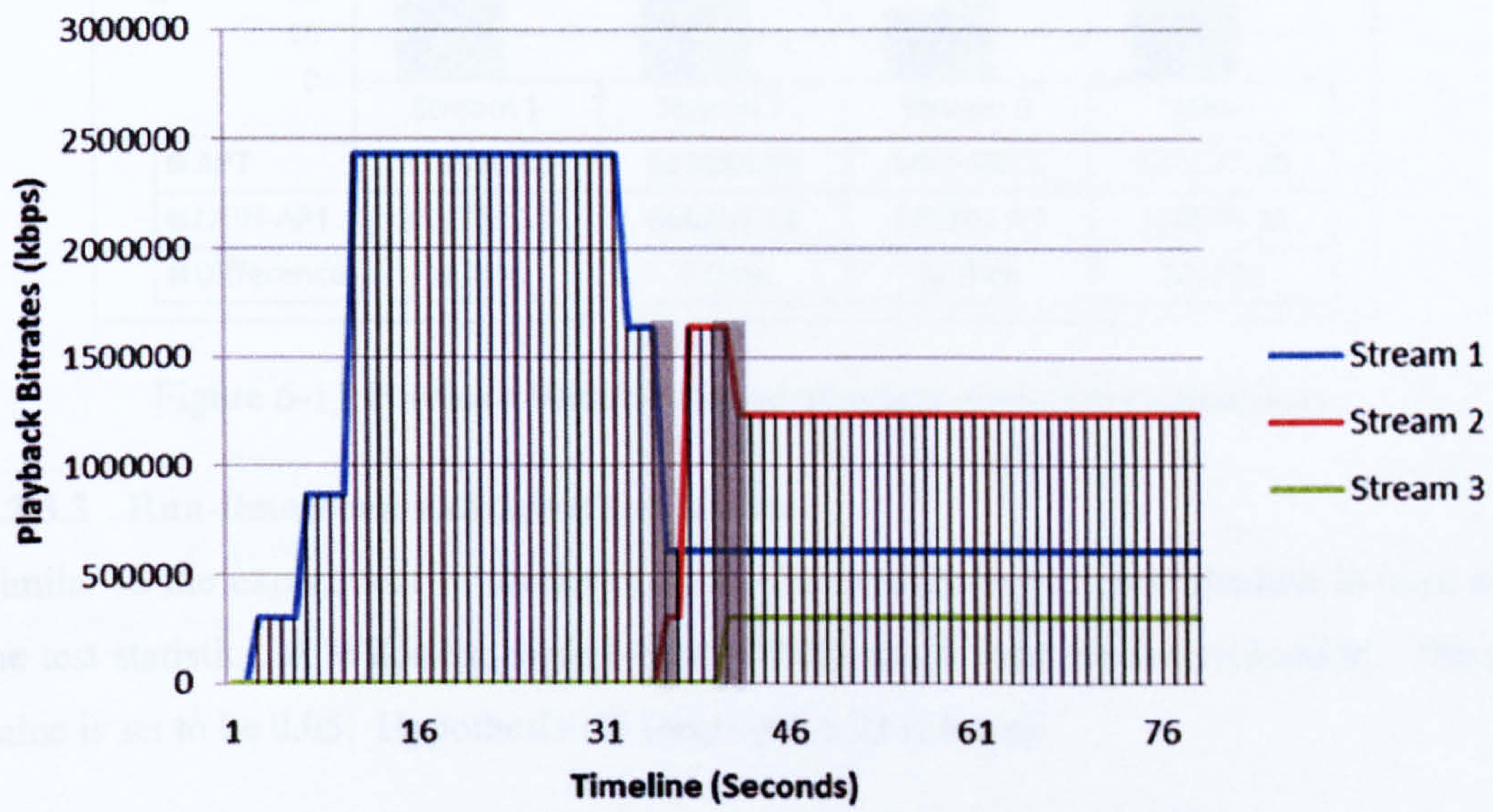
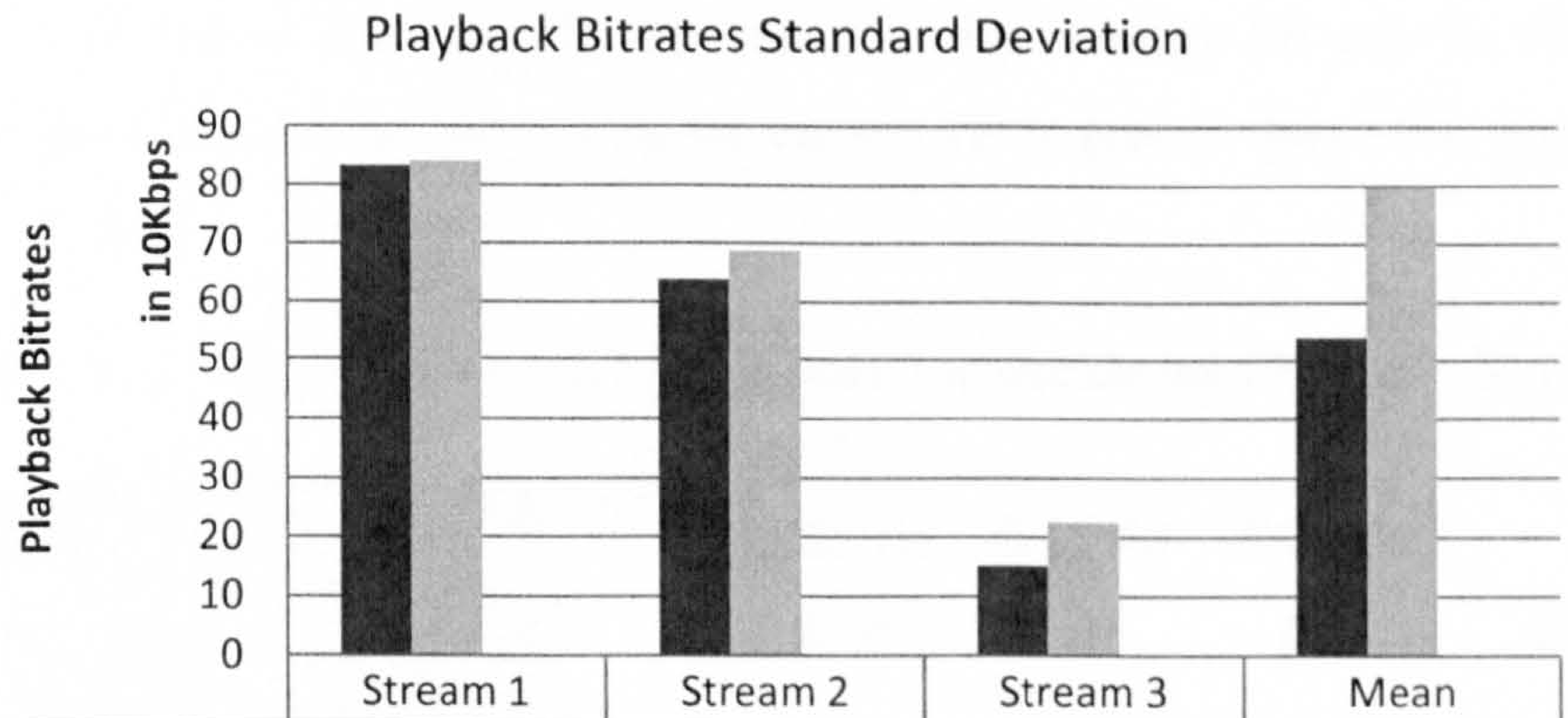


Figure 6-10 Bitrates changes with multi-stream adaptation

After 45s, the system with multi-stream adaptation has stable playback bitrates for all three streams. The system without multi-stream adaptation encounters another three bandwidth competition instances.

In Figure 6-11, a comparison of playback bitrates standard deviation is illustrated. In this experiment, despite the fact that the former existing streams have a higher standard deviation than latter incoming ones, the system with multi-stream adaptation tends to always have a lower standard deviation than the system without multi-stream adaptation. The standard deviation difference between the two systems also tends to increase with respect to each newly incoming stream. This suggests that the bandwidth competition problem becomes more severe when there are more streams.

To sum up this experiment, **the results suggest that for the proposed multi-stream adaptation mechanism, the bandwidth competition problem can be effectively alleviated (i.e. 54.84% on average in this experiment) and that more stable playback bitrates can be supported.**



	Stream 1	Stream 2	Stream 3	Mean
APT	833576.46	637988.70	150946.38	540837.18
NON-APT	839733.33	686267.64	225102.99	798576.26
Difference	0.73%	7.04%	32.94%	32.27%

Figure 6-11 Playback bitrates change standard deviation comparison

6.2.3.3 Run-time Personalisation Evaluation

Similar to the experiment in section 6.2.2.2, the prediction precision median is used as the test statistics. A Wilcoxon signed-rank (WSR) test is used for the evaluation. The p value is set to be 0.05. Hypothesis H5 (section 5.6.2) is tested.

Table 6-21 Hypothesis testing with PP median, p=0.05

User ID	Prediction Precision Generate by Personalisation Mechanism								
1	0.75	0.75	0.75	0.75	0.75	**0.25** ($p<0.05$)	0.50 ($p>0.05$)	0.50 ($p>0.05$)	0.75 ($p>0.05$)
2	0.25	0.25	0.25	0.25	0.25	0.50 ($p>0.05$)	0.50 ($p>0.05$)	0.50 ($p>0.05$)	0.75 ($p>0.05$)
3	0.75	0.75	0.75	0.75	0.75	0.75 ($p>0.05$)	0.75 ($p>0.05$)	0.75 ($p>0.05$)	**0.50** ($p<0.05$)
4	1.00	1.00	1.00	1.00	1.00	**0.75** ($p<0.05$)	**0.50** ($p<0.05$)	**0.50** ($p>0.05$)	0.75 ($p>0.05$)
5	0.75	0.75	0.75	0.75	0.75	**0.25** ($p<0.05$)	**0.25** ($p<0.05$)	**0.25** ($p<0.05$)	**0.25** ($p<0.05$)
6	0.25	0.25	0.25	0.25	0.25	0.25 ($p>0.05$)	0.75 ($p>0.05$)	0.50 ($p>0.05$)	0.50 ($p>0.05$)

In Table 6-21, the bold numbers indicate when the hypothesis is rejected. The average hypothesis rejection rate is 56.25%[13]. The test enables the system to adjust itself to perform the personalisation for the next camera switch session based upon the following rule:

Rule 6.2.3.3-1: If the hypothesis is rejected, the system can replace the current prediction precision (PP) with a dummy PP, the value of this dummy PP can be obtained by reducing the current PP by 15% given the current PP is greater than 15%, otherwise the value will be 0 .

Table 6-22 Applying rule 6.2.3.3-1 to user 5 to decide the PP threshold value

User 5	0.75	0.75	0.75	0.75	0.75	0.25 To 0.10 ($p<0.05$)	0.25 To 0.10 ($p<0.05$)	0.25 To 0.10 ($p<0.05$)	0.25 ($p>0.05$)

By applying the Rule 6.2.3.3-1 to user 5, the hypothesis rejection rate will be reduced from 100% to 75% as shown in Table 6-22. Then the hypothesis rejection rate is declined to 50%. **This result suggests that a rule based PP threshold is able to reduce the hypothesis test rejection rate.**

[13] 4 users with 16 PP values, 9 PP values out of 16 had been rejected by the hypothesis test.

Table 6-23 Hypothesis testing with PP median produced by random approach, p=0.05

User ID	Prediction Precision Generate by Personalisation Mechanism								
1	0.75	0.75	0.75	0.75	0.75	0.25 ($p>0.05$)	0.50 ($p>0.05$)	0.50 ($p>0.05$)	0.75 ($p>0.05$)
2	0.25	0.25	0.25	0.25	0.25	0.50 ($p>0.05$)	0.50 ($p>0.05$)	0.50 ($p>0.05$)	0.75 ($p>0.05$)
3	0.75	0.75	0.75	0.75	0.75	0.75 ($p>0.05$)	0.75 ($p>0.05$)	0.75 ($p>0.05$)	0.50 ($p>0.05$)
4	1.00	1.00	1.00	1.00	1.00	0.75 ($p>0.05$)	0.50 ($p>0.05$)	0.50 ($p>0.05$)	0.75 ($p>0.05$)
5	0.75	0.75	0.75	0.75	0.75	0.25 ($p>0.05$)	0.25 ($p>0.05$)	0.25 ($p>0.05$)	0.25 ($p>0.05$)
6	0.25	0.25	0.25	0.25	0.25	0.25 ($p>0.05$)	0.75 ($p>0.05$)	0.50 ($p>0.05$)	0.50 ($p>0.05$)

Table 6-23 shows the results to test hypothesis H6. The results indicate there is no rejection of the hypothesis. **This suggests that the personalised system has a higher PP median than a system that lacks personalisation support.**

The application of the test results enables the system to compare automatic camera switching, with or without personalisation using the following rule:

Rule 6.2.3.3-2: If the hypothesis is rejected continuously, e.g. 3 sequential times, the system can disable the automatic camera switching. The automatic camera switching activates only when the hypothesis is not rejected for the previous multi-angle viewing session.

Personalisation mechanism enables better automatic camera switching when applying this rule at run-time (Table 6-23). Note when the rule is further applied to Table 6-21, the personalization effectively applied to user 4 and 5 at use session 5 given the threshold value for hypothesis continuous rejection is 3. This is illustrated in Table 6-24.

Table 6-24 Applying Rule 6.2.3.3-2 to decide to activated auto camera switching for users 4 and 5 in Table 6-21

User 4	1.00	1.00	1.00	1.00	1.00	**0.75** ($p<0.05$)	**0.50** ($p<0.05$)	**0.50** ($p>0.05$)	0.75 ($p>0.05$)
User 5	0.75	0.75	0.75	0.75	0.75	**0.25** ($p<0.05$)	**0.25** ($p<0.05$)	**0.25** ($p<0.05$)	**0.25** ($p<0.05$)

As shown in the table above, the personalisation will stop in the next use session when the PP reaches 0.5 and will reactivate after the PP value reaches 0.75 for user 4. For user 5, the personalisation will be inactivated when PP value reaches 0.25. Hence, **the group recommender system is able to adjust its performance using an appropriate rule.**

6.2.3.4 Post-Operation Evaluation

The consistency test is based upon the hypothesis H7. In this experiment, a normality test is initially conducted with Jarque-Bera test, and F-test. For non-normally distributed sample data, a WSR test will be used to test whether median of this data is greater or equal than the average median of the consistent sample data. If this test is not rejected, the tested non-normally distributed sample data will join the consistent sample data group.

Based upon the data in Table 6-20, four users' (user 1, 2, 4 and 6) data are normally distributed. The consistency rate based upon F-test of these users is 100% with an average PP median of 0.56. The PP median value of the WSR test (left-tailed test)[14] for two groups of non-normally distributed user data (user 3 and 5) also indicates that the hypothesis is not rejected when the given median value of 56.25% with p value of 0.05. **As a result, it can be concluded that the consistency rate of personalisation across these six users is 4/6, i.e. 66.67% which suggests the personalisation is consistent.**

The hypothesis for the scalability test is H8. The test is applied to the six users. A Spearman's rank correlation coefficient test is used. Table 6-25 shows the results of the correlation of number of uses and the prediction precision of each user.

Table 6-25 Correlation coefficient of number of uses and prediction precision

User ID	Correlation Coefficient ρ	Correlation Coefficient Significance (left tail, p=0.05)
1	0.16	1.00 (>0.05)
2	0.89	1.00 (>0.05)
3	**-0.61**	1.00 (>0.05)
4	**-0.48**	1.00 (>0.05)
5	**-0.61**	1.00 (>0.05)
6	-0.63	1.00 (>0.05)

[14] The test is based upon the null hypothesis that the PP median of normally distributed data groups is less or equal to the testing user's median value

The results indicate that half of the users have positive correlation coefficient values. If defines the correlations as: [0.5~1]: strong positive correlation, [-1~-0.5] strong negative correlation), (0~0.5) weak positive correlation, and (-0.5~0) weak negative correlation. 33.3% of the users (i.e. user 2 and 6) will have a strong positive correlation coefficient and another 33.3% of the users (i.e. user 3 and 5) will have strong negative correlation coefficient, one user (i.e. user 4) will have weak negative correlation coefficient and one user (i.e. user 1) will have a weak positive correlation coefficient. **Hence, it can be concluded that the hypothesis is not rejected and that personalisation is scalable.**

6.2.4 Personalised Selective Target Zooming

The video ZUI capability of the developed Web based system and the quality of the personalisation are evaluated for this user task.

6.2.4.1 Experimental Data Set

The video ZUI components are tested in a laboratory setting in which the network bandwidth is strictly controlled. Two types of sports event were considered: a single player game and a multi-player game. The athletic long jump and 400m video clips from Aviva European Trials and UK Championships 2008 are chosen as experimental video content. Videos with an original resolution of 640 X 480 and a frame rate of 25fps were encoded at different bitrates versions. The bitrates used are: 230Kbps, 305Kbps, 403Kbps, 543Kbps, 708Kbps, 937Kbps, 1241Kbps and 1644Kbps. Video streams were looped to simulate a live mode. Note that high definition content was not available from real video events.

Six users including both expert users and novice users participated in the personalisation experiments. In the experiments, 60% of users' raw historical data in the usage part of the user profile is used as the training data. Each user was asked to use the system for more than at least 15 zooming sessions respectively (i.e. a session starts by zooming in and finishes by zooming out) for each of the two different sports events. The personalisation mechanism was executed before and after user's manual zooming so that accumulated prediction precision value can be obtained after each zooming session. The true prediction here is defined as the predicted zooming cluster actually being zoomed. The system also randomly generated a zooming centroid based upon front end video screen resolution representing the results for the system without personalisation. The system PP value is generated based upon the fact that whether the random centroid falls into the actually cluster representing the region of interest for zooming. The first 11 zooming sessions' data of each user are used as some users had only performed 11 zoom-in actions.

Table 6-26 and Table 6-27 show the data retrieved from long jump event. Table 6-28 and Table 6-29 show the data retrieved from the 400m event. The first 11 zooming sessions' data of each user are used as some users had only performed 11 zoom-in actions.

Table 6-26 Accumulated prediction precision values of 6 users in 11 personalised zooming sessions for the long jump event

Zooming Session	User 1	User 2	User 3	User 4	User 5	User 6
1	0.00	1.00	1.00	1.00	1.00	0.00
2	0.50	1.00	0.50	1.00	1.00	0.00
3	0.67	0.67	0.33	1.00	0.67	0.33
4	0.75	0.75	0.50	1.00	0.75	0.25
5	0.60	0.80	0.60	1.00	0.80	0.40
6	0.50	0.67	0.67	0.83	0.67	0.50
7	0.57	0.57	0.71	0.86	0.71	0.57
8	0.63	0.63	0.63	0.75	0.63	0.63
9	0.67	0.67	0.56	0.78	0.67	0.67
10	0.60	0.70	0.60	0.80	0.70	0.60
11	0.64	0.73	0.64	0.82	0.73	0.64

Table 6-27 Accumulated prediction precision values of 6 users in 11 non-personalised zooming sessions for the long jump event

Zooming Session	User 1	User 2	User 3	User 4	User 5	User 6
1	0.00	1.00	0.00	0.00	0.00	0.00
2	0.00	0.50	0.00	0.50	0.00	0.00
3	0.00	0.33	0.00	0.33	0.33	0.00
4	0.25	0.50	0.00	0.25	0.25	0.25
5	0.20	0.40	0.20	0.40	0.40	0.20
6	0.17	0.33	0.33	0.33	0.33	0.17
7	0.14	0.43	0.29	0.29	0.29	0.29
8	0.25	0.37	0.25	0.25	0.36	0.25
9	0.22	0.33	0.22	0.22	0.33	0.22
10	0.30	0.40	0.20	0.20	0.30	0.20
11	0.27	0.36	0.18	0.27	0.27	0.18

Table 6-28 Accumulated prediction precision values of 6 users in 11 personalised zooming sessions for the 400m event

Zooming Session	User 1	User 2	User 3	User 4	User 5	User 6
1	1.00	0.00	0.00	1.00	1.00	0.00
2	0.50	0.00	0.50	0.50	0.50	0.50
3	0.33	0.33	0.67	0.67	0.33	0.33
4	0.50	0.25	0.75	0.75	0.25	0.50
5	0.60	0.40	0.60	0.80	0.20	0.40
6	0.50	0.33	0.50	0.67	0.33	0.50
7	0.57	0.43	0.43	0.71	0.43	0.57
8	0.63	0.38	0.50	0.75	0.50	0.50
9	0.56	0.44	0.56	0.67	0.56	0.56
10	0.50	0.50	0.60	0.70	0.60	0.60
11	0.55	0.55	0.64	0.73	0.64	0.55

Table 6-29 Accumulated prediction precision values of 6 users in 11 non-personalised zooming sessions for the 400m event

Zooming Session	User 1	User 2	User 3	User 4	User 5	User 6
1	1.00	0.00	0.00	0.00	0.00	0.00
2	1.00	0.00	0.00	0.00	0.00	0.00
3	0.65	0.00	0.00	0.00	0.00	0.00
4	0.50	0.00	0.25	0.00	0.00	0.00
5	0.40	0.00	0.20	0.00	0.00	0.00
6	0.33	0.00	0.17	0.00	0.17	0.17
7	0.42	0.14	0.14	0.00	0.13	0.13
8	0.38	0.13	0.13	0.00	0.25	0.25
9	0.33	0.11	0.22	0.00	0.22	0.22
10	0.30	0.10	0.20	0.00	0.20	0.20
11	0.27	0.09	0.18	0.00	0.18	0.27

6.2.4.2 Personalisation Task Interface Evaluation

Two experiments were conducted to evaluate video ZUI performance focusing on the zooming video quality adaptation performance and to test the effectiveness of zooming with respect to the time-shift playback and zooming animation duration. For the first Zooming Video Quality Adaptation Performance experiment, the long jump event was the focus. A terminal with screen resolution of 1600(W) X 900(H) hosted the player and Google chrome was used as the Web browser. The system was launched with a video screen of 951 X 500, and a zooming factor was 1.3. The zooming animation duration was set to be 300 milliseconds.

Figure 6-12 shows the adaptation time latency of the video stream bitrate adaptation. The adaptation time latency is the time difference between each zooming animation finish and any the change in the resulting video bitrate. The collected adaptation time latency samples, for 22 sample zooming actions, were analysed further using WSR tests[15] (left-tailed test and right-tailed test). The test result revealed that the population median time latency ranged from 79.90 to 92.60 milliseconds ($p \geq 0.05$). **The results suggest that given a frame rate of 25fps, video bitrate adaptation seems to occur within three sequential video frames after zooming.**

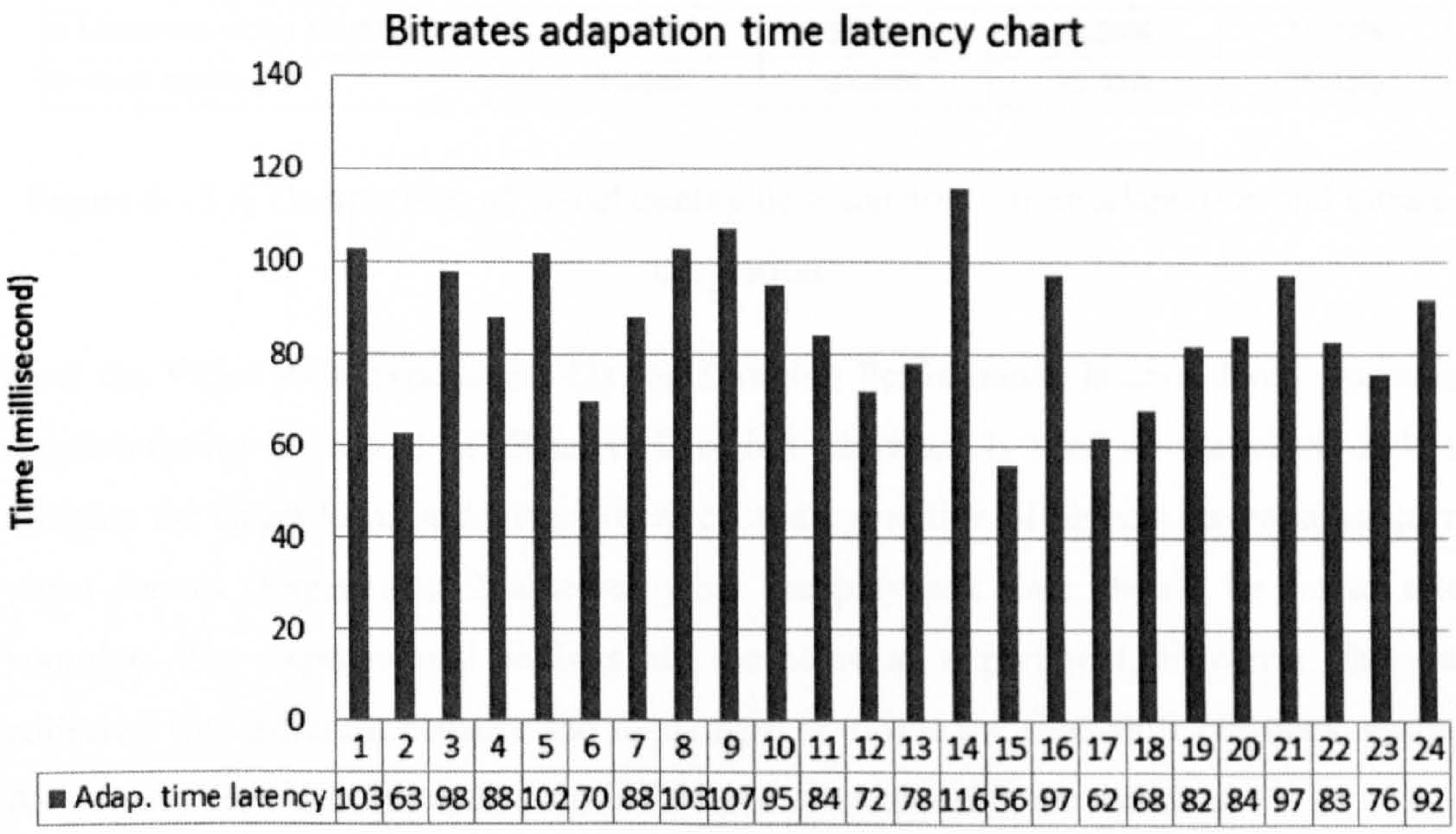
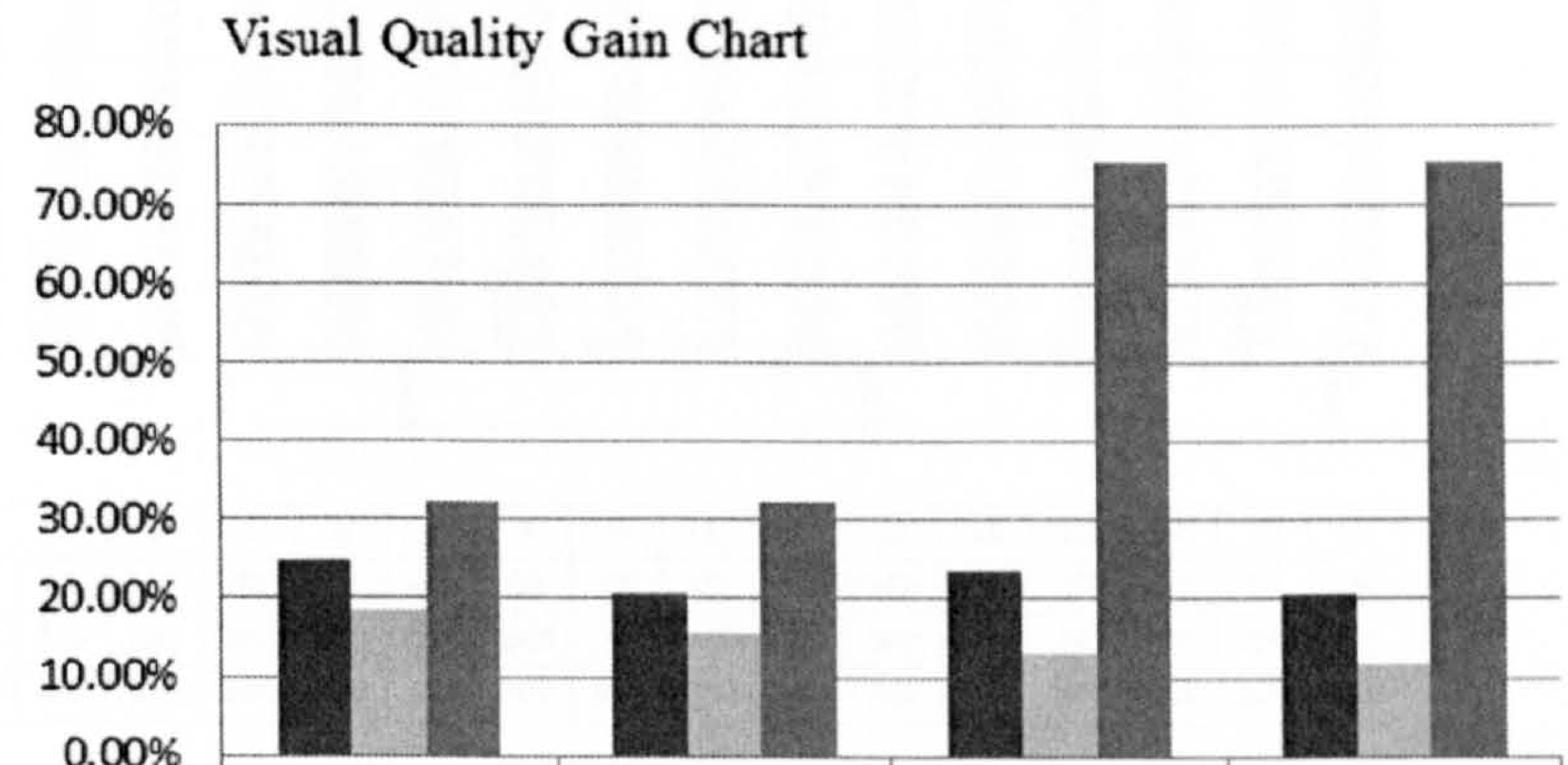


Figure 6-12 Time latency between zooming events and bitrates changes

To further illustrate the video quality gain produced using bitrate adaptation, a comparison of the bitrates-to-video height ratio between non-bitrate adaptation (NA) zooming and bitrate adaptation (A) zooming is given. The initial video screen height is 500 pixels and the associated bitrate is 937Kbps. For non-adaptation zooming, the bitrates will always be 937Kbps. When adaptation zooming is used with 500 and 600 pixel height screens, the adapted bitrates is 1241Kbps. For 700 and 800 pixel height screens, the adapted bitrate is 1644Kbps. The results in Figure 6-13 indicate that the

[15] Left/Right -tailed test is based upon the null hypothesis that the median of obtained time cost is greater/less than a given value. Here, the test is used to obtain the maximum value given value with p=0.05.

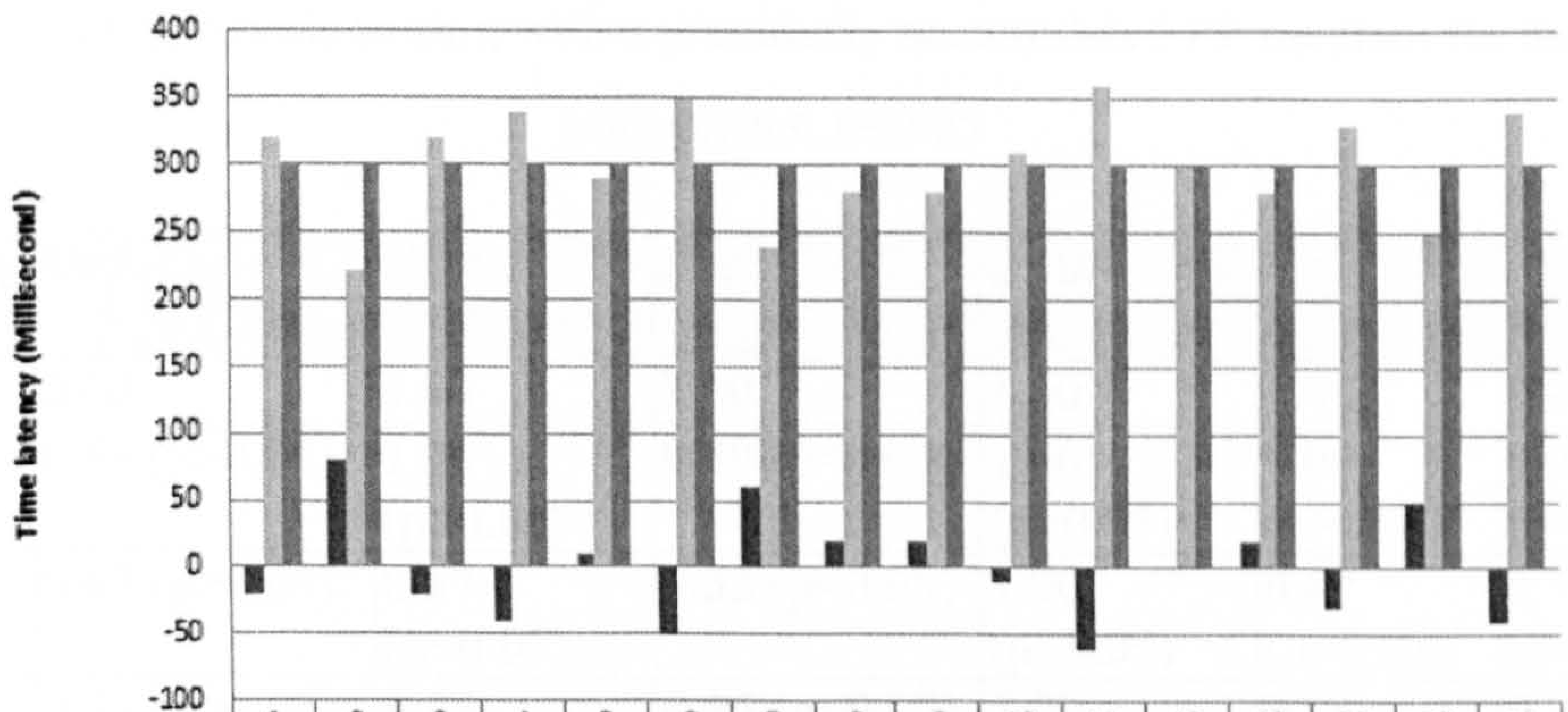
visual quality gain (i.e. (A-NA)/NA) reaches 75.45% when zooming using a higher resolution video screen.



	500	600	700	800
bitrates-to-video height ratio (A)	24.82%	20.68%	23.49%	20.55%
bitrates-to-video height ratio (NA)	18.74%	15.62%	13.39%	11.71%
visual quality gain	32.44%	32.44%	75.45%	75.45%

Figure 6-13 A Comparison of visual quality between non-bitrate adaptation and bitrate adaptation

Next the Video ZUI Experiment (2) for Zooming Performance is considered. Accurate target locating is critical in ZUI. A time-shift playback is used in the video ZUI to mitigate the target location shift problem caused by motion of objects across subsequent video frames. Experiment 2 assesses what the playback time should be set to after zooming. The experimental settings are the same as experiment. However, the data collected was different because the focus here is to test the time-shift playback setting. As a result, the video time positions before and after zooming are recorded.



	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Time Latency	-20	80	-20	-40	10	-50	60	20	20	-10	-60	0	20	-30	50	-40
Actual time	320	220	320	340	290	350	240	280	280	310	360	300	280	330	250	340
Ideal time	300	300	300	300	300	300	300	300	300	300	300	300	300	300	300	300

Figure 6-14 Time latency for time-shift playback

The data collected is shown in Figure 6-14. The data shows that there is a time latency associated with each time-shift playback. The time latency in this experiment is defined as the difference between the ideal shifted time (i.e. 300 milliseconds) and the actual shifted time. A WSR (both left-tailed and right-tailed) test16 reveals that the population median time latency is between -20.02 and 20.07 milliseconds ($p \geq 0.05$). Hence, the use of **time-shift playbacks enable users to keep track of a particular image frame containing a zooming target.**

6.2.4.3 Run-time Personalisation Evaluation

Similar to the experiment in section 6.2.3.3, the prediction precision median is used as the test statistic to test hypothesis H5 (section 5.6.2). A Wilcoxon signed-rank (WSR) test is used for the evaluation, and the p value is set to be 0.05.

[16] Left/Right -tailed test is based upon a null hypothesis when the median of the obtained latency is greater/less than a given value. Here, the test is used to obtain the maximum given value with p=0.05.

Table 6-30 Hypothesis testing with a previously accumulated PP median, for the long jump event, p=0.05

Zooming Session	User 1	User 2	User 3	User 4	User 5	User 6
1	0.00	1.00	1.00	1.00	1.00	0.00
2	0.50 ($p>0.05$)	1.00 ($p>0.05$)	**0.50** ($p<0.05$)	1.00 ($p>0.05$)	1.00 ($p>0.05$)	0.00 ($p>0.05$)
3	0.67 ($p>0.05$)	**0.67** ($p<0.05$)	**0.33** ($p<0.05$)	1.00 ($p>0.05$)	**0.67** ($p<0.05$)	0.33 ($p>0.05$)
4	0.75 ($p>0.05$)	**0.75** ($p<0.05$)	**0.50** ($p<0.05$)	1.00 ($p>0.05$)	**0.75** ($p<0.05$)	0.25 ($p>0.05$)
5	0.60 ($p>0.05$)	**0.80** ($p<0.05$)	0.60 ($p>0.05$)	1.00 ($p>0.05$)	**0.80** ($p<0.05$)	0.40 ($p>0.05$)
6	0.50 ($p>0.05$)	**0.67** ($p<0.05$)	0.67 ($p>0.05$)	**0.83** ($p<0.05$)	**0.67** ($p<0.05$)	0.50 ($p>0.05$)
7	0.57 ($p>0.05$)	**0.57** ($p<0.05$)	0.71 ($p>0.05$)	**0.86** ($p<0.05$)	**0.71** ($p<0.05$)	0.57 ($p>0.05$)
8	0.63 ($p>0.05$)	**0.63** ($p<0.05$)	0.63 ($p>0.05$)	**0.75** ($p<0.05$)	**0.63** ($p<0.05$)	0.63 ($p>0.05$)
9	0.67 ($p>0.05$)	**0.67** ($p<0.05$)	**0.56** ($p<0.05$)	**0.78** ($p<0.05$)	**0.67** ($p<0.05$)	0.67 ($p>0.05$)
10	0.60 ($p>0.05$)	**0.70** ($p<0.05$)	0.60 ($p>0.05$)	**0.80** ($p<0.05$)	**0.70** ($p<0.05$)	0.60 ($p>0.05$)
11	0.64 ($p>0.05$)	0.73 ($p>0.05$)	0.64 ($p>0.05$)	**0.82** ($p<0.05$)	0.73 ($p>0.05$)	0.64 ($p>0.05$)

In Table 6-30, the bold numbers indicate when the hypothesis is rejected. The average hypothesis rejection rate is 70%[17].

[17] 4 users with 40 PP values, 28 PP values out of 40 had been rejected by the hypothesis test.

Table 6-31 Hypothesis testing with previous accumulated PP median, p=0.05 (400m)

Zooming Session	User 1	User 2	User 3	User 4	User 5	User 6
1	1.00	0.00	0.00	1.00	1.00	0.00
2	**0.50** ($p<0.05$)	0.00 ($p>0.05$)	0.50 ($p>0.05$)	**0.50** ($p<0.05$)	**0.50** ($p<0.05$)	0.50 ($p>0.05$)
3	**0.33** ($p<0.05$)	0.33 ($p>0.05$)	0.67 ($p>0.05$)	**0.67** ($p<0.05$)	**0.33** ($p<0.05$)	0.33 ($p>0.05$)
4	**0.50** ($p<0.05$)	0.25 ($p>0.05$)	0.75 ($p>0.05$)	0.75 ($p>0.05$)	**0.25** ($p<0.05$)	0.50 ($p>0.05$)
5	0.60 ($p>0.05$)	0.40 ($p>0.05$)	0.60 ($p>0.05$)	0.80 ($p>0.05$)	**0.20** ($p<0.05$)	0.40 ($p>0.05$)
6	**0.50** ($p<0.05$)	0.33 ($p>0.05$)	0.50 ($p>0.05$)	**0.67** ($p<0.05$)	**0.33** ($p<0.05$)	0.50 ($p>0.05$)
7	0.57 ($p>0.05$)	0.43 ($p>0.05$)	0.43 ($p>0.05$)	**0.71** ($p<0.05$)	0.43 ($p>0.05$)	0.57 ($p>0.05$)
8	0.63 ($p>0.05$)	0.38 ($p>0.05$)	0.50 ($p>0.05$)	0.75 ($p>0.05$)	0.50 ($p>0.05$)	0.50 ($p>0.05$)
9	0.56 ($p>0.05$)	0.44 ($p>0.05$)	0.56 ($p>0.05$)	**0.67** ($p<0.05$)	0.56 ($p>0.05$)	0.56 ($p>0.05$)
10	**0.50** ($p<0.05$)	0.50 ($p>0.05$)	0.60 ($p>0.05$)	**0.70** ($p<0.05$)	0.60 ($p>0.05$)	0.60 ($p>0.05$)
11	0.55 ($p>0.05$)	0.55 ($p>0.05$)	0.64 ($p>0.05$)	0.73 ($p>0.05$)	0.64 ($p>0.05$)	0.55 ($p>0.05$)

In Table 6-31, the average hypothesis rejection rate is 53.33%. Due to the fact that the rejection could happen when the prediction is actually right (e.g. in the 10th zooming session of user 2, Table 6-30,), the application of the test results obtained can be used to allow the system to adjust itself to overcome this problem and to reduce the rejection rate at the same time using the rule:

Rule 6.2.4.3-1: If the hypothesis is rejected, the system can replace the current accumulated prediction precision (PP) with by reducing the current PP variable by 15% given the PP is greater than 15%, otherwise the value will be 0.

Rule 6.2.4.3-1 is applied to users 2, 3, 4 and 5 (Table 6-30) as shown in Table 6-32. For users 2 and 5, hypothesis has been accepted in sessions of 5, 9 and 10. For user 4, hypothesis has been accepted in sessions 11. The hypothesis rejection rate is also reduced from 70% to 47.5%. When the rule is applied to users 1, 3 and 4 (Table 6-31), hypothesis has been accepted in session 7, 9 and 10 for user 4 and session 5 of user 5. The average hypothesis rejection rate is further reduced from 53.33%% to 36.67%

Table 6-32 Applying Rule 6.2.4.3-1 to decide the threshold PP value on user 2, 3, 4 and 5 (Table 6-30)

Zooming Session	User 2	User 3	User 4	User 5
1	1.00	1.00	1.00	1.00
2	1.00 (p>0.05)	**0.50 to 0.35** (p<0.05)	1.00 (p>0.05)	1.00 (p>0.05)
3	**0.67 to 0.52** (p<0.05)	**0.33 to 0.18** (p<0.05)	1.00 (p>0.05)	**0.67 to 0.52** (p<0.05)
4	**0.75 to 0.60** (p<0.05)	**0.50 to 0. 35** (p<0.05)	1.00 (p>0.05)	**0.75 to 0.60** (p<0.05)
5	*0.80 (p>0.05)*	0.60 (p>0.05)	1.00 (p>0.05)	*0.80 (p>0.05)*
6	**0.67 to 0.52** (p<0.05)	0.67 (p>0.05)	**0.83 to 0.68** (p<0.05)	**0.67 to 0.52** (p<0.05)
7	**0.57 to 0.42** (p<0.05)	0.71 (p>0.05)	**0.86 to 0.71**(p<0.05)	**0.71 to 0.56** (p<0.05)
8	**0.63 to 0.48** (p<0.05)	0.63 (p>0.05)	**0.75 to 0.60** (p<0.05)	**0.63 to 0.48** (p<0.05)
9	*0.67 (p>0.05)*	**0.56 to 0.41** (p<0.05)	**0.78 to 0.63** (p<0.05)	*0.67 (p>0.05)*
10	*0.70 (p>0.05)*	0.60 (p>0.05)	**0.80 to 0.65** (p<0.05)	*0.70 (p>0.05)*
11	0.73 (p>0.05)	0.64 (p>0.05)	*0.82 (p>0.05)*	0.73 (p>0.05)

Table 6-33 Applying Rule 6.2.4.3-1 to decide the threshold PP value on user 1, 4 and 5(Table 6-31)

Zooming Session	User 1	User 4	User 5
1	1.00	1.00	1.00
2	**0.50 to 0.35** (p<0.05)	**0.50 to 0.35** (p<0.05)	**0.50 to 0.35** (p<0.05)
3	**0.33 to 0.18** (p<0.05)	**0.67 to 0.52** (p<0.05)	**0.33 to 0.18** (p<0.05)
4	**0.50 to 0.35** (p<0.05)	0.75 (p>0.05)	**0.25 to 0.10** (p<0.05)
5	0.60 (p>0.05)	0.80 (p>0.05)	**0.20 to 0.05** (p<0.05)
6	*0.50 (p<0.05)*	**0.67 to 0.52** (p<0.05)	*0.33* (p>0.05)
7	0.57 (p>0.05)	*0.71* (p>0.05)	0.43 (p>0.05)
8	0.63 (p>0.05)	0.75 (p>0.05)	0.50 (p>0.05)
9	0.56 (p>0.05)	*0.67 (p>0.05)*	0.56 (p>0.05)
10	**.50 to 0.35** (p<0.05)	*0.70* (p>0.05)	0.60 (p>0.05)
11	0.55 (p>0.05)	0.73 (p>0.05)	0.64 (p>0.05)

By referring to the test input data given in The first 11 zooming sessions' data of each user are used as some users had only performed 11 zoom-in actions.

Table 6-26 and Table 6-27, Table 6-34 shows the hypothesis test results for hypothesis H5 for the long jump. The results indicate there is no rejection of the hypothesis. **This suggests that the personalised system has a higher PP median than the system without personalisation support.** With reference to the data in Table 6-28 and Table 6-29, Table 6-35 shows the hypothesis testing results for 400m. This indicates that for the first six zooming sessions of user 1, a lower PP median value results compared to the system without personalisation.

Table 6-34 Hypothesis testing with a previously accumulated PP median produced by random approach, for the Long jump event p=0.05.

Zooming Session	User 1	User 2	User 3	User 4	User 5	User 6
1	0.00	1.00	1.00	1.00	1.00	0.00
2	0.50 (p>0.05)	1.00 (p>0.05)	0.50 (p>0.05)	1.00 (p>0.05)	1.00 (p>0.05)	0.00 (p>0.05)
3	0.67 (p>0.05)	0.67 (p>0.05)	0.33 (p>0.05)	1.00 (p>0.05)	0.67 (p>0.05)	0.33 (p>0.05)
4	0.75 (p>0.05)	0.75 (p>0.05)	0.50 (p>0.05)	1.00 (p>0.05)	0.75 (p>0.05)	0.25 (p>0.05)
5	0.60 (p>0.05)	0.80 (p>0.05)	0.60 (p>0.05)	1.00 (p>0.05)	0.80 (p>0.05)	0.40 (p>0.05)
6	0.50 (p>0.05)	0.67 (p>0.05)	0.67 (p>0.05)	0.83 (p>0.05)	0.67 (p>0.05)	0.50 (p>0.05)
7	0.57 (p>0.05)	0.57 (p>0.05)	0.71 (p>0.05)	0.86 (p>0.05)	0.71 (p>0.05)	0.57 (p>0.05)
8	0.63 (p>0.05)	0.63 (p>0.05)	0.63 (p>0.05)	0.75 (p>0.05)	0.63 (p>0.05)	0.63 (p>0.05)
9	0.67 (p>0.05)	0.67 (p>0.05)	0.56 (p>0.05)	0.78 (p>0.05)	0.67 (p>0.05)	0.67 (p>0.05)
10	0.60 (p>0.05)	0.70 (p>0.05)	0.60 (p>0.05)	0.80 (p>0.05)	0.70 (p>0.05)	0.60 (p>0.05)
11	0.64 (p>0.05)	0.73 (p>0.05)	0.64 (p>0.05)	0.82 (p>0.05)	0.73 (p>0.05)	0.64 (p>0.05)

Table 6-35 Hypothesis Testing with a previous accumulated PP median produced by a system without personalisation for the 400M event, p=0.05

Zooming Session	User 1	User 2	User 3	User 4	User 5	User 6
1	1.00	0.00	0.00	1.00	1.00	0.00
2	**0.50** (p<0.05)	0.00 (p>0.05)	0.50 (p>0.05)	0.50 (p>0.05)	0.50 (p>0.05)	0.50 (p>0.05)
3	**0.33** (p<0.05)	0.33 (p>0.05)	0.67 (p>0.05)	0.67 (p>0.05)	0.33 (p>0.05)	0.33 (p>0.05)
4	**0.50** (p<0.05)	0.25 (p>0.05)	0.75 (p>0.05)	0.75 (p>0.05)	0.25 (p>0.05)	0.50 (p>0.05)
5	**0.60** (p<0.05)	0.40 (p>0.05)	0.60 (p>0.05)	0.80 (p>0.05)	0.20 (p>0.05)	0.40 (p>0.05)
6	**0.50** (p<0.05)	0.33 (p>0.05)	0.50 (p>0.05)	0.67 (p>0.05)	0.33 (p>0.05)	0.50 (p>0.05)
7	**0.57** (p<0.05)	0.43 (p>0.05)	0.43 (p>0.05)	0.71 (p>0.05)	0.43 (p>0.05)	0.57 (p>0.05)
8	0.63 (p>0.05)	0.38 (p>0.05)	0.50 (p>0.05)	0.75 (p>0.05)	0.50 (p>0.05)	0.50 (p>0.05)
9	0.56 (p>0.05)	0.44 (p>0.05)	0.56 (p>0.05)	0.67 (p>0.05)	0.56 (p>0.05)	0.56 (p>0.05)
10	0.50 (p>0.05)	0.50 (p>0.05)	0.60 (p>0.05)	0.70 (p>0.05)	0.60 (p>0.05)	0.60 (p>0.05)
11	0.55 (p>0.05)	0.55 (p>0.05)	0.64 (p>0.05)	0.73 (p>0.05)	0.64 (p>0.05)	0.55 (p>0.05)

The application of the results also allows the system to adjust its performance by activating or deactivating the personalised zooming using the rule:

Rule 6.2.4.3-2: If the hypothesis is rejected continuously, e.g. 3 sequential times, the system can disable the personalised zooming. Personalised zooming activates only when the hypothesis is not rejected in the previous zooming session.

By applying this rule to Table 6-35 at run-time, suggests that in this example, the system should stop the personalisation in session 5 and restart it in session 9, (Table 6-36). When this rule is applied to the Table 6-32 and Table 6-33, further false predictions can be avoided.

Table 6-36 Applying rule 6.2.4.3-2 to decide activation of personalised zooming on User 1 (Table 6-35)

Zooming Session	User 1
1	1.00
2	**0.50** (p<0.05)
3	**0.33** (p<0.05)
4	**0.50** (p<0.05)
5	**0.60** (p<0.05)
6	**0.50** (p<0.05)
7	**0.57** (p<0.05)
8	0.63 (p>0.05)
9	0.56 (p>0.05)
10	0.50 (p>0.05)
11	0.55 (p>0.05)

As seen from table above, the personalisation will stop in use session 4 after PP value reaches 0.5, and will be reactivated in session 9 after PP value reaches 0.63 in use session 8. **As a result, it can be concluded that the group recommender system is able to adjust its performance using an appropriate rule.**

6.2.4.4 Post-Operation Evaluation

The consistency test is based upon the hypothesis H7. Based on the data given in The first 11 zooming sessions' data of each user are used as some users had only performed 11 zoom-in actions.

Table 6-26, two users' data (user 1, and 5) are not normally distributed. The consistency rate based upon an F-test for normally distributed user data is 50% with an average accumulated PP median of 0.68. Hence, based upon hypothesis testing, it can be concluded that the consistency rate of personalisation across these six users is 2/6, i.e. 33.33%. The accumulated PP median value WSR test (left-tailed test)[18] for of non-normally distributed user data (user 1 and 5) also indicate that the hypothesis is not rejected indicated by the given median value of 0.68 with p = 0.05. Hence, the overall consistency rate is 66.67%. For the data in Table 6-27, data of the users 1, 3 and 6 are not normally distributed. The consistency rate for the rest of the user data is 66.67% with an average accumulated PP median of 0.61. The consistency rate of personalisation of these six users therefore is 2/6, i.e. 33.33%, which is the same as for the long jump video

[18] The test is based upon the null hypothesis that the PP median of normally distributed data groups is less or equal to the testing user's median value

scenario. **The accumulated PP median value WSR test shows that one of the three non-normally distributed users' data set is not rejected with p=0.05, which suggests an overall consistency rate for all users is not less than 50%. Hence, personalisation is consistent.**

For the scalability test, the hypothesis H8 is tested using the Spearman's rank correlation coefficient test on 6 users. Table 6-37 shows the results of the correlation of the number of uses and the prediction precision of each user.

Table 6-37 Correlation coefficient for a number of uses and prediction precision

User ID	Correlation Coefficient ρ	Correlation Coefficient Significance (left tail, p=0.05)
1	0.05	1.00 (>0.05)
2	0.93	1.00 (>0.05)
3	0.21	1.00 (>0.05)
4	**-0.07**	1.00 (>0.05)
5	0.29	1.00 (>0.05)
6	0.79	1.00 (>0.05)

If defines the correlations as: [0.5~1]: strong positive correlation, [-1~-0.5] strong negative correlation), (0~0.5) weak positive correlation, and (-0.5~0) weak negative correlation. The results indicate that 5/6, i.e. 83.33% of the users have positive correlation coefficient values. Using the correlations defined in 6.2.3.4, 33.3% of the users (users 2 and 6) will have a strong positive correlation coefficient and another 50% of the users (users 1, 3 and 5) have a weak positive correlation coefficient. One user (user 4) has a weak negative correlation coefficient. **Therefore, it can be concluded that the hypothesis is not rejected and that personalisation is scalable.**

6.2.5 Time-Shift Viewing

The evaluation in this section assesses the performance of the incidents highlighting and objects highlighting.

6.2.5.1 Experimental Data Set

Two athletic events including women's long jump (single player game) and men's 400m (multi-player game) were used to evaluate this task. Both these events are pre-recorded from a single camera view.

6.2.5.2 Incidents Highlighting Evaluation

The evaluation of this incident driven task is based upon a set of predefined incidents. In this experiment, incidents for two video clips are studied beforehand. The duration of a competitive 400m event is about 1 minute. Each triple long jump attempt is less than half that time. In Table 6-38, predefined incidents and highlighted incidents are listed with the corresponding timeline positions for two sample sports events. Bold timeline positions represent the highlighted incidents.

Table 6-38 Incidents highlighting test results

Testing Video Content	Pre-identified sports incident (Description)	Highlighted critical timelines shortly before sports incidents occur (Description)
#1 men's 400 m (1 minute 8 seconds)	1. 00:00:10.173 (run) 2. 00:00:21.909 (1st corner) 3. 00:00:22.168 (overtake) 4. 00:00:28.889 (overtake) 5. 00:00:31.371 (overtake) 6. 00:00:35.145 (overtake) 7. 00:00:37.679 (2nd corner) 8. 00:00:42.901 (3rd corner) 9. 00:00:49.881 (last 100 meter close up) 10. 00:01:00.428 (1st athlete crosses the finish line) 11. 00:01:03.737 (last athlete crosses the finish line)	**1. 00:00:09.181 (before run)** **2. 00:00:35.084 (before overtake)** 3. **00:00:35.629 (before second corner)** **4. 00:00:38.659 (before third corner)** 5. 00:00:38.758 (before third corner) **6. 00:00:43.273 (before last 100 meters)** 7. 00:00:43.901 (before last 100 meters) 8. 00:00:48.444 (before last 100 meters) **9. 00:00:50.523 (before first athlete crosses the finish line)** 10. 00:00:50.624 (before first one cross finish line) 11. 00:00:58.774 (before first one cross finish line) 12. 00:00:58.853 (before first one cross finish line) **13. 00:01:00.922 (before last athlete crosses the finish line)**
#2 women's long jump (23.40 seconds)	1. 00:00:04.320 (ready) 2. 00:00:13.393 (run) 3. 00:00:18.867 (touch take off board) 4. 00:00:20.080 (last jump) 5. 00:00:20.720 (in the pit) 6. 00:00:22.720 (walk out the pit)	**1. 00:00:01.914 (before ready)** 2. 00:00:02.958 (before ready) 3. 00:00:03.513 (before ready) **4. 00:00:09.748 (before run)** 5. 00:00:10.451 (before run) **6. 00:00:15.278 (before touch take off board)** **7. 00:00:21.147 (before walk out the pit)**

The recall and precision for the highlighted incidents are defined using equations in Section 5.6.2.2. The incidents recall for both sports events are 63.64% and 66.67% according to Table 6-38. The incident highlighting precision is evaluated in terms of a 1 second, 2 second and 3 second lead time. The precision results shown in Table 6-39 indicate that a higher precision can be obtained with a shorter lead time. **The results suggest that the system is able to highlight the incidents in a one second lead time with a high precision.**

Table 6-39 Incident highlighting precision in terms of lead time

Testing Video	3 S Precision	2 S Precision	1 S Precision
#1	6 /13 = 46.15%	10/13 = 76.92%	12/13 = 92.31%
#2	2/7 = 28.57%	5/7 = 71.43%	6/7 = 85.71%

6.2.5.3 Object Highlighting Evaluation

Figure 6-15 and Figure 6-16 demonstrate two scenarios with the images on the left-hand side showing the histograms and the centroids of the highlighting objects. The right-hand side image shows the corresponding visual annotations. **Objects of interest can be highlighted in both a single object scenario (upper pair of images) and a multi-object scenario (lower pair of images).**

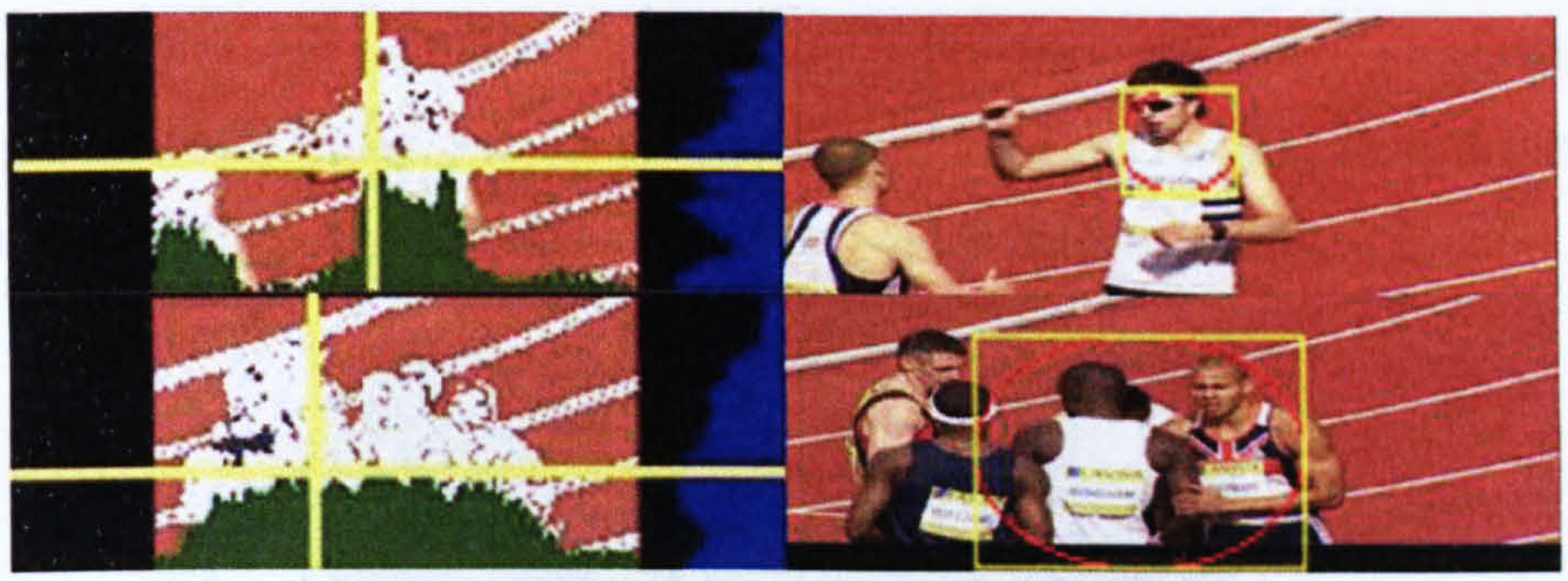


Figure 6-15 400m Event objects highlighting histograms vs. visual annotation

Figure 6-16 Long jump event objects highlighting histograms versus visual annotation

6.2.5.4 Time-shift Viewing Experience

Here the application of the results obtained from the evaluation is demonstrated. Figure 6-17 shows a plot of all the incidents, highlights, objects highlighting/replay (OH/RP) against a timeline from the start of a 400 M event. In addition, the replay periods are also highlighted on the timeline. Replays start from the nearest highlights point and should be at least 10 seconds before the OH/RP point. 10 seconds is the default replay duration.

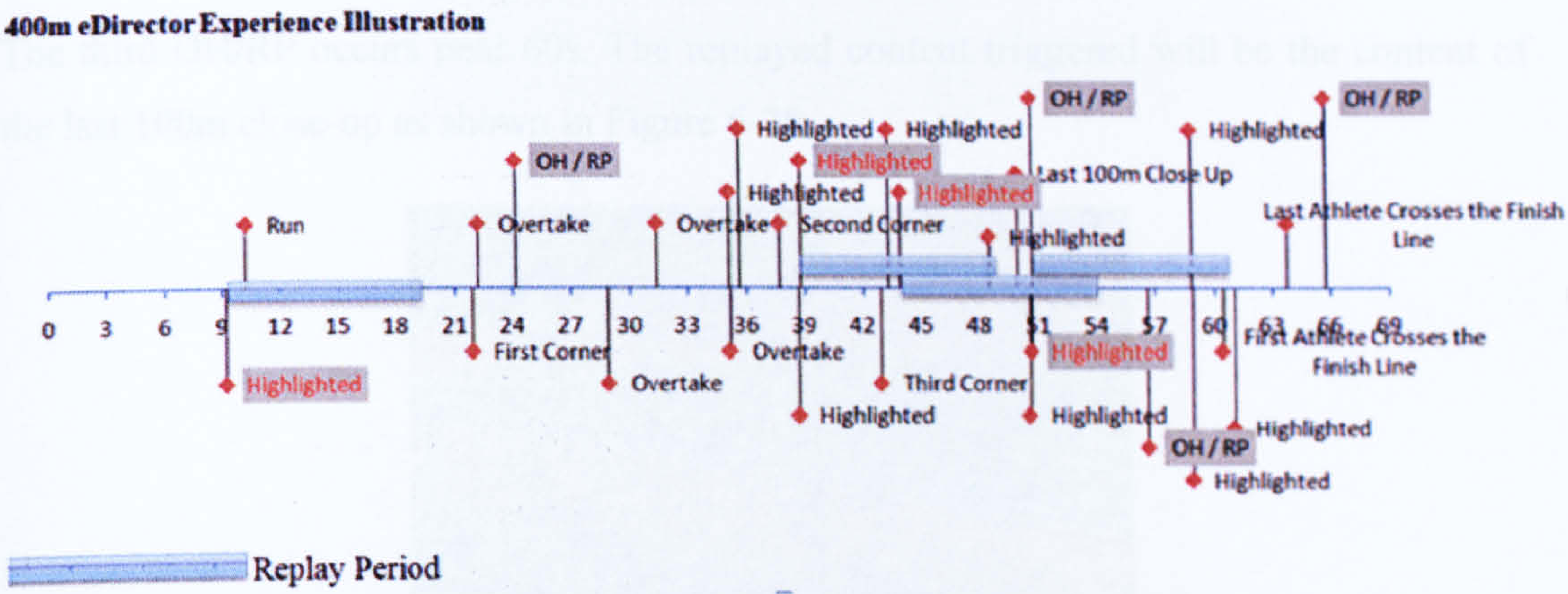


Figure 6-17 Incidents and highlights are plotted onto the event timeline (time unit = second)

The first OH/RP occurs around 24s. The system can perform either object highlighting or remind the user of a replay. For the case of replay, incidents running from the start line will be replayed. Figure 6-18 shows a first frame of the replayed scene.

Figure 6-18 Start running replay scene (first frame)

The second OH/RP occurs around 51s. If the replay is accepted by the user, the event highlights around 39s will be replayed, and scenes when athletes turn the third corner will be replayed. One of these scenes is shown in Figure 6-19.

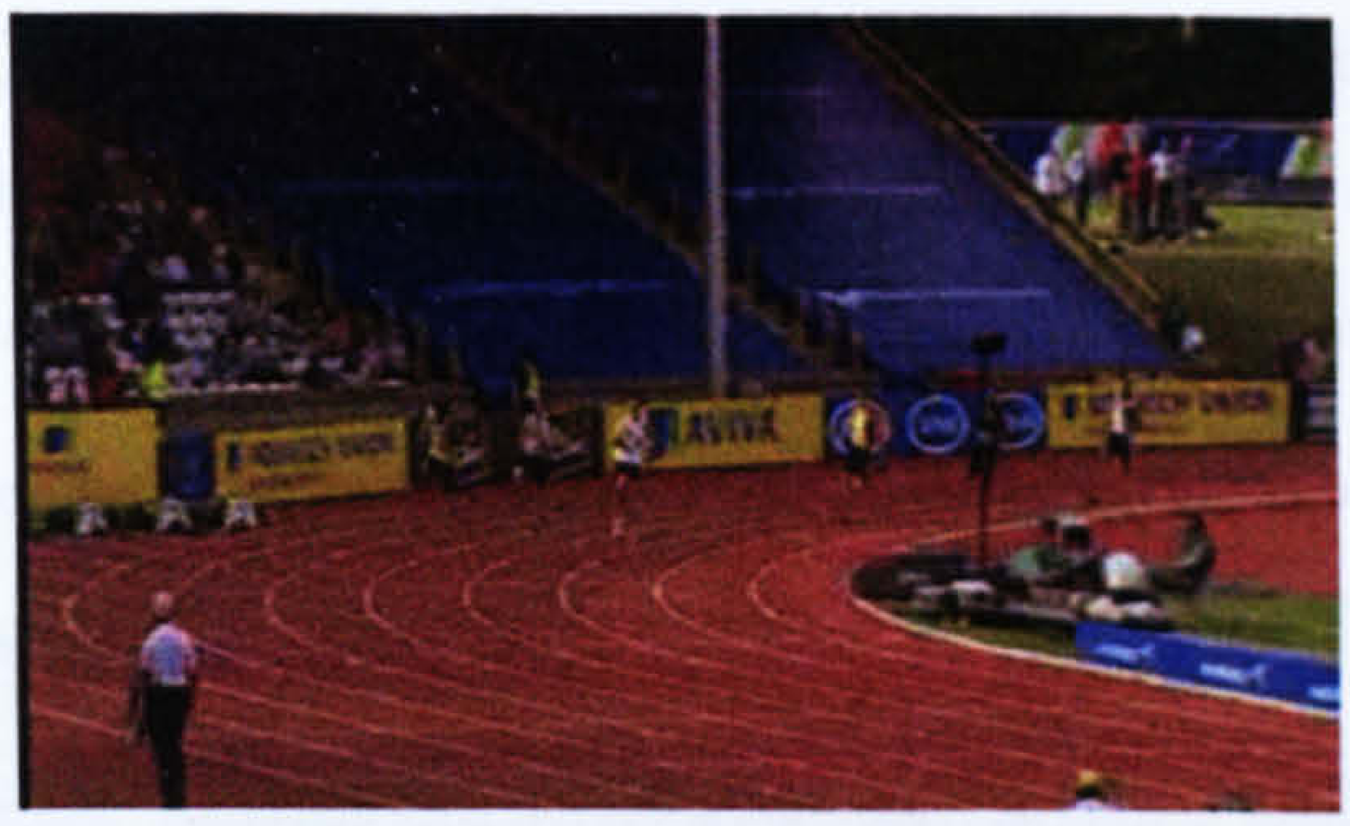


Figure 6-19 Third corner turning replay scene

The third OH/RP occurs near 60s. The replayed content triggered will be the content of the last 100m close-up as shown in Figure 6-20.

Figure 6-20 Last 100m close up replay scene

The fourth OH/RP occurs near 66s. The replay will cover whole process of the last 100m. Figure 6-21 shows the last captured replay frame showing the first athlete crossing the finish line.

Figure 6-21 First Athlete crosses the finish line replay scene (last frame)

Object highlighting and replays can also trigger zooming and camera switching based upon detected incidents (Figure 6-22, Figure 6-23, Figure 6-24 and Figure 6-25). The first OH/RP point gives a far view of athletes. Therefore zooming in on a highlighted area could allow views more clearly display athletes. As an alternative, a second camera could also bring to users a detailed view of the athletes. Similarly, this also applies to the second and third OH/RP points which view the athletes from a medium distance. For the fourth OH/RP point, a short distance view is given. In addition to object highlighting, the replay and alternative camera view perhaps give viewers a better view.

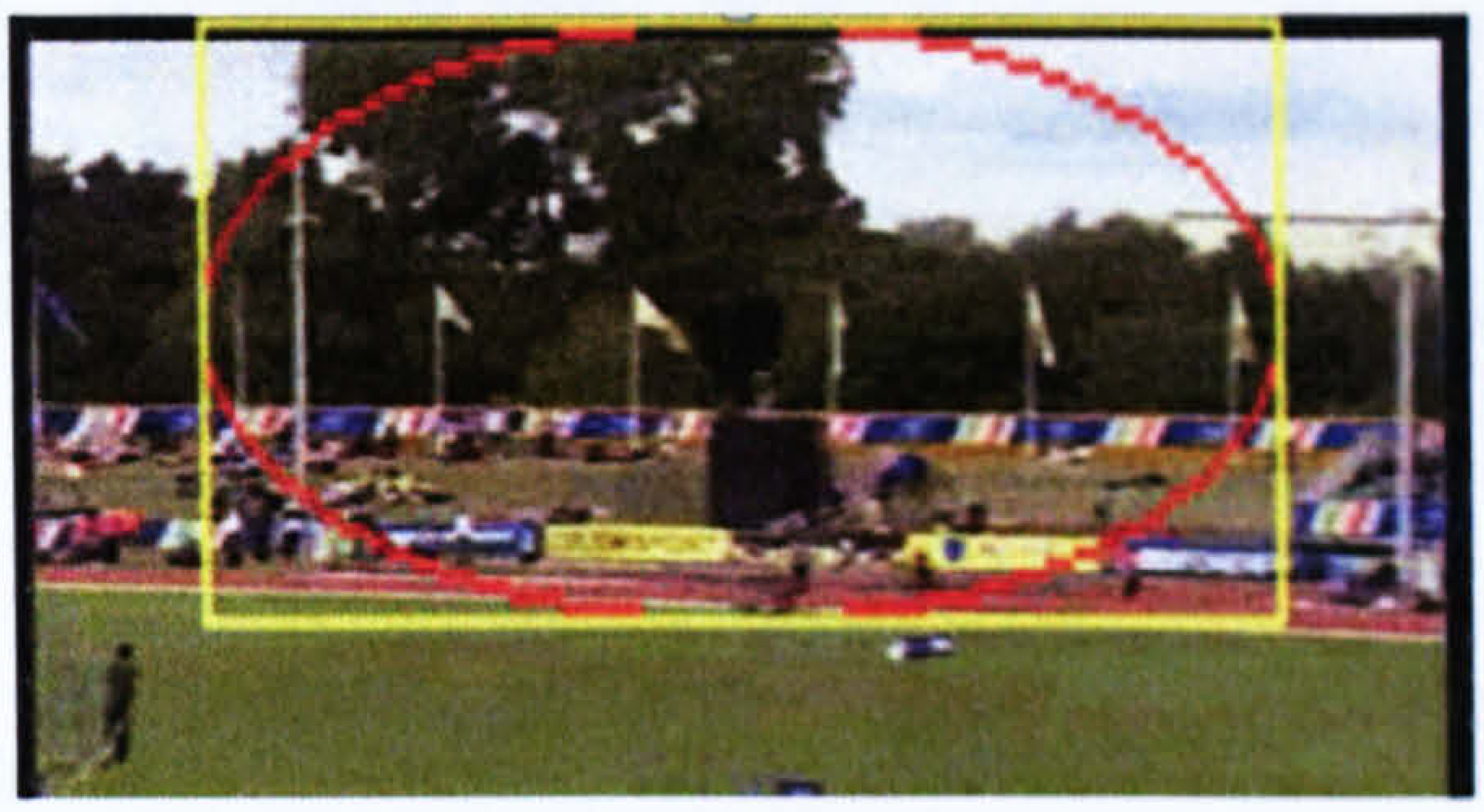


Figure 6-22 First OH/RP point

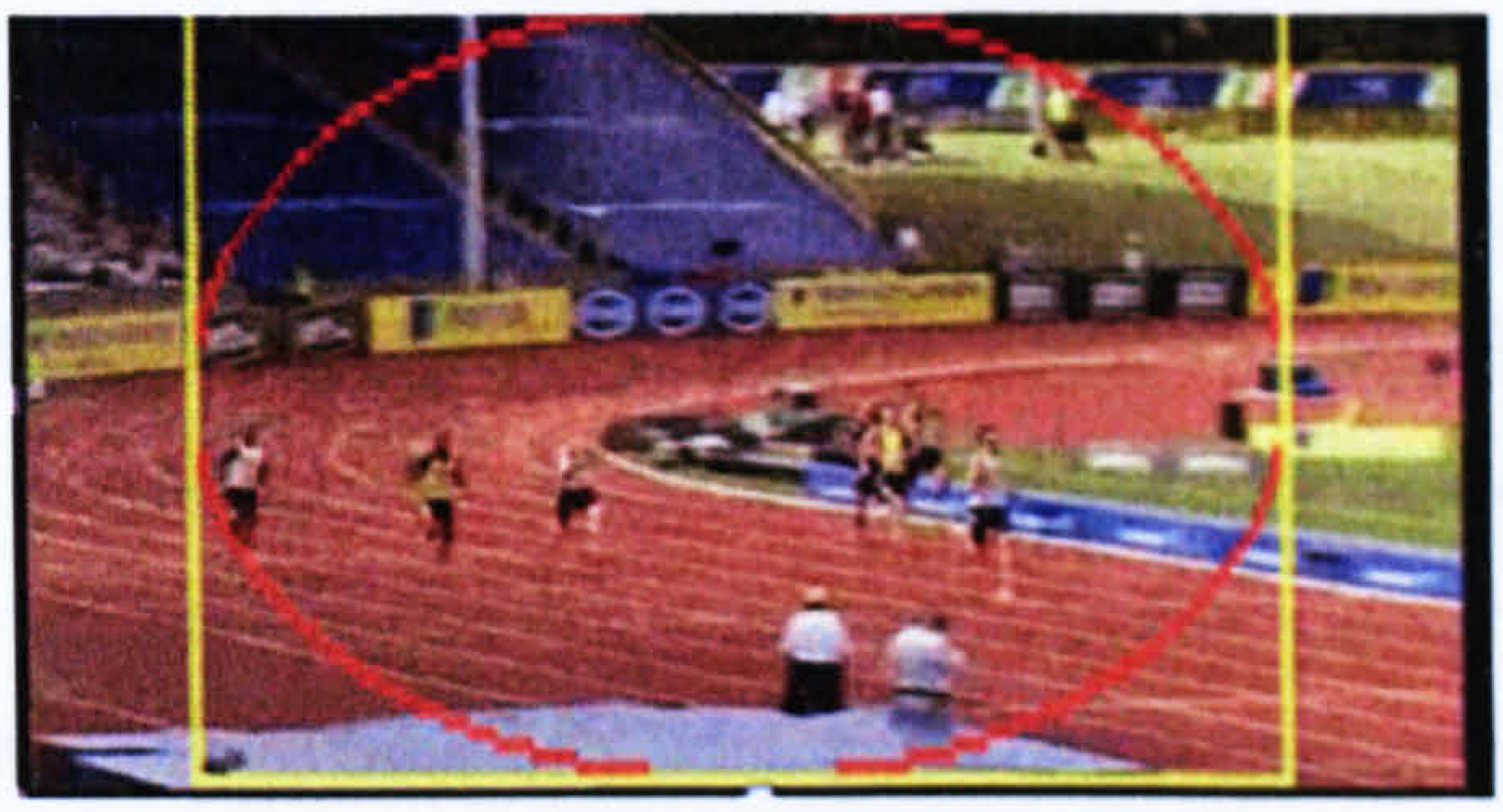


Figure 6-23 Second OH/RP point

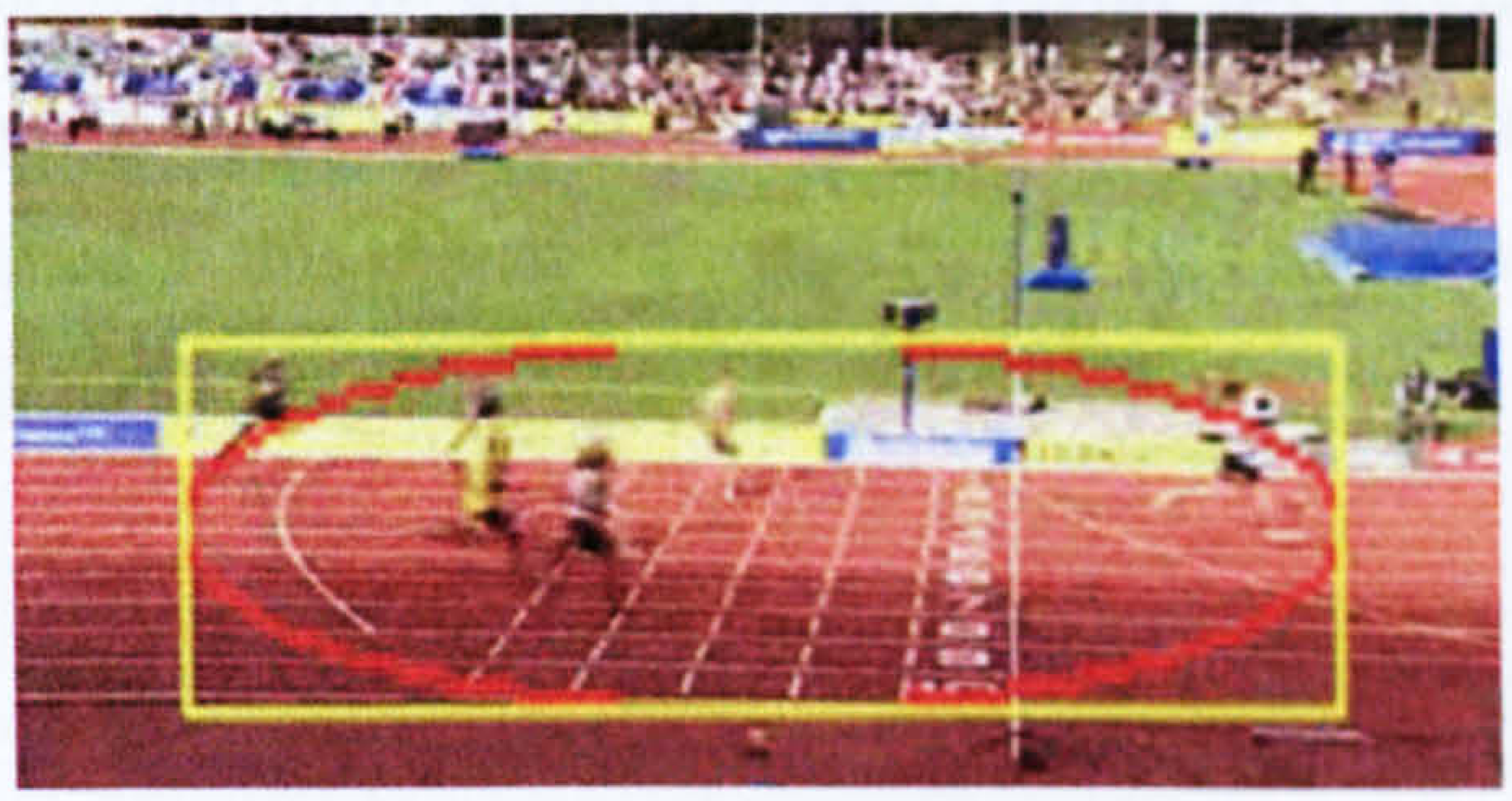


Figure 6-24 Third OH/RP point

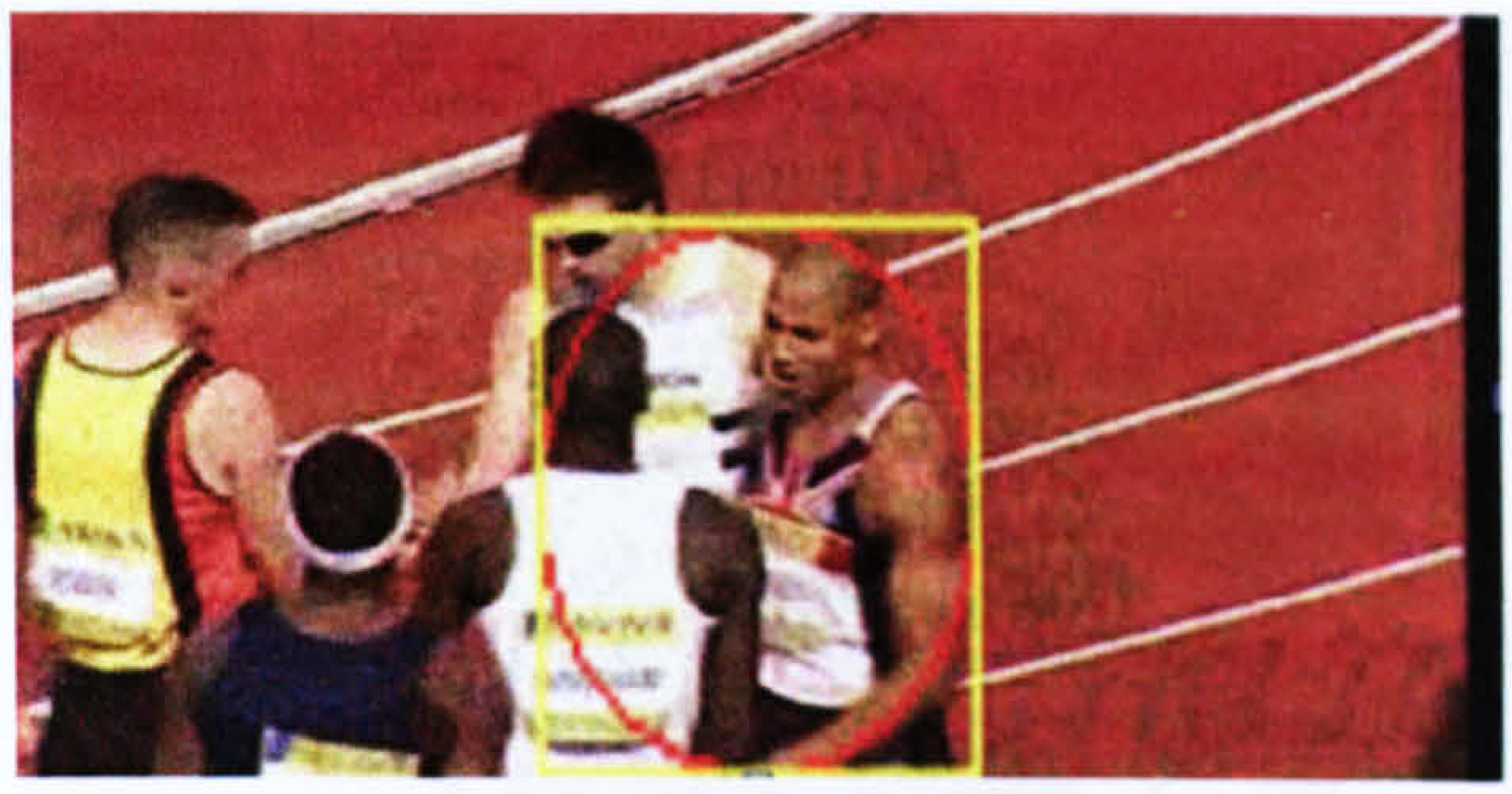


Figure 6-25 Fourth OH/RP point

6.3 Usability Evaluation

The usability evaluation was based upon the user feedback and a post-trial online questionnaire in relation to a pre-trial questionnaire. The pre-trial questionnaire includes three questions about the age, gender and Internet TV use frequency of users. The post-trial questionnaire consists of 21 questions which was designed using usability experts from both Queen Mary University of London (QMUL) and British Broadcasting Corporation (BBC) for use in the My eDirector 2012 project. Users could access this questionnaire either via the link on the player page or from the link sent by a system administrator by email. In this trial, fifty users participated (the same group of users as in the experiment described in section 6.2.1.1).

Figure 6-26 shows the main UI of the trial system. Figure 6-27 shows the questionnaire page. The A-V player's name used in this author's trials is called 'eDirector'. The usability results obtained in this trial also formed part of this author's contributions to the EU Project My eDirector 2012.

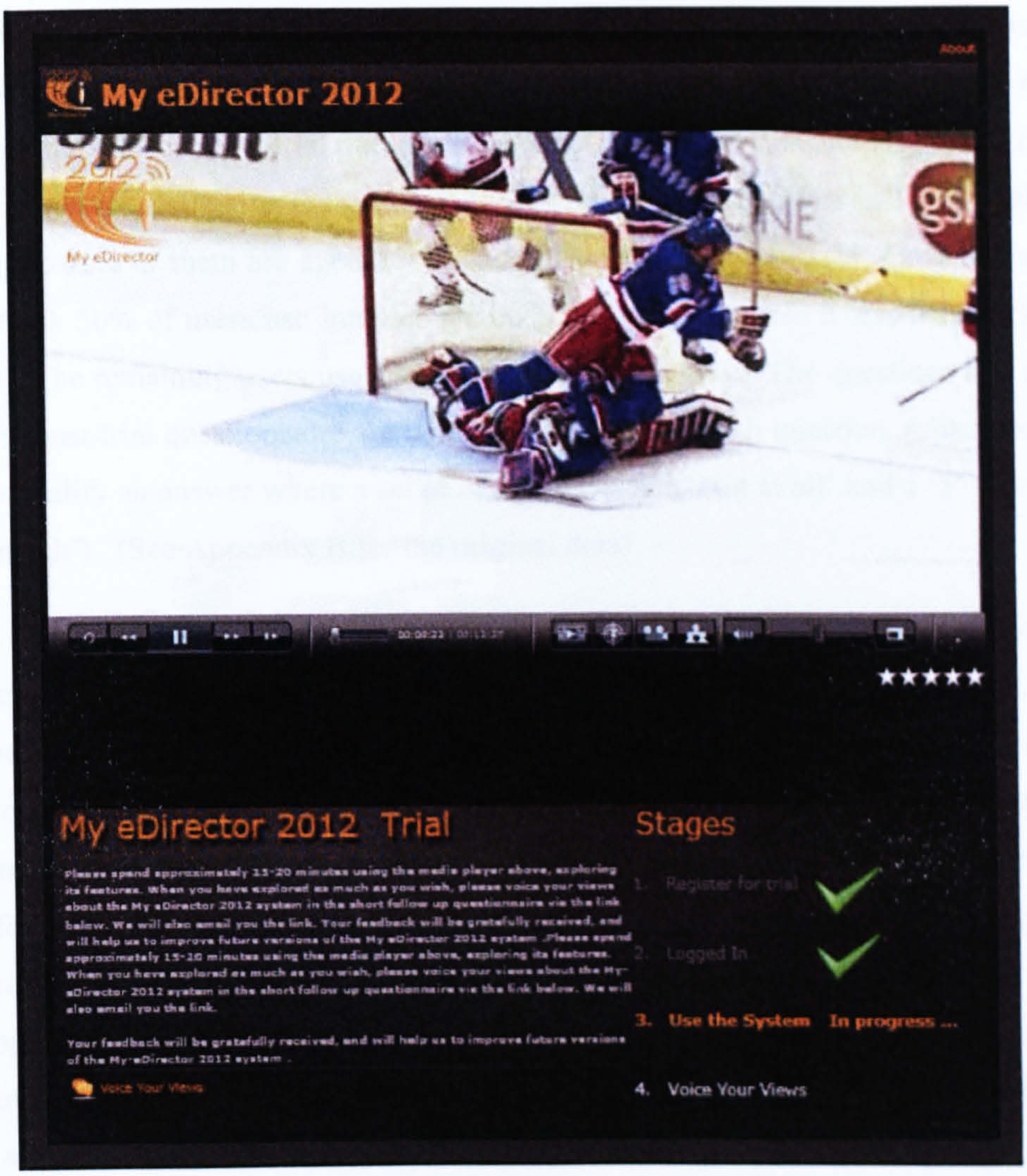


Figure 6-26 Player main page

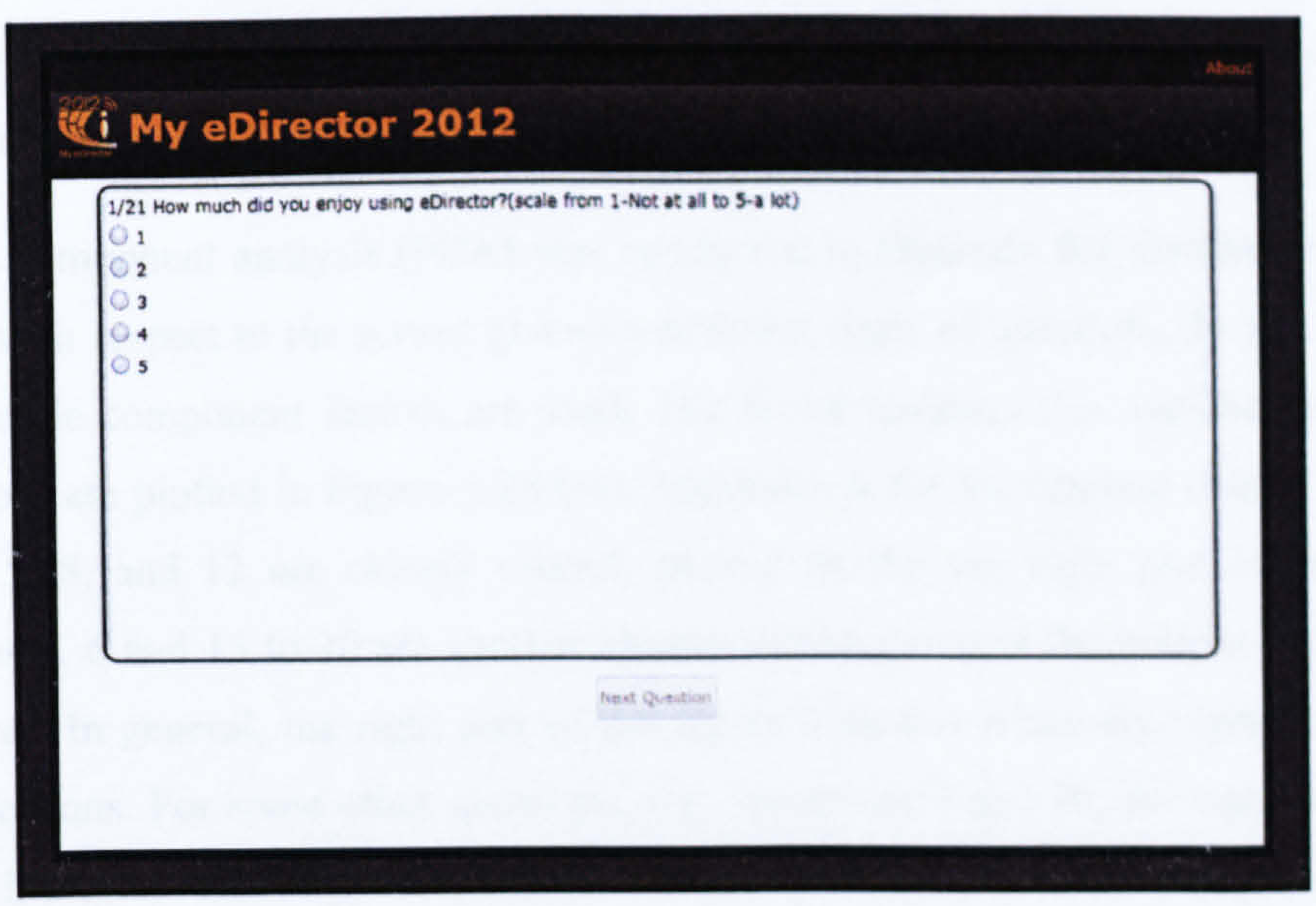


Figure 6-27 Questionnaire page

Fifty users completed the pre-trial questionnaire; 26 users finished the post-trial questionnaire. In order to correlate to the pre-trial and post-trial question answers, the 26 users that finished the post-trial questionnaire were studied here. According to the results from pre-trial questionnaire, the gender distribution for these 26 users is 27% female and 73% male. 62% of them are aged 26-35 and 35% of users aged 18-25. Only one user is aged 36-50. 50% of users use Internet TV daily. 15% of users use it less frequently (i.e. weekly). The remaining users use Internet TV less frequently. The questions and results from the post-trial questionnaire are described below. For each question, a likert scale is used to qualify an answer where a score of '1' represents 'Not at all' and a '5' represents 'Very much'). (See Appendix B for the original data).

In this post-trial questionnaire, questions 1 to 5, users were asked about the overall experience of using the player. On average 84% (score 4 and 5) of the users had a positive feeling. 92% of users would most likely to recommend the system to others. Question 6, 7 and 8 asked about users' attitude towards the features of camera switching, recommendation, and directing. The response shows that most users were interested in these features. Question 9 and 10 asked about the learnability of the player. As expected, most users felt it was easy to learn and use. Question 11 and 12 asked about users' viewing experience in terms of video quality and smoothness. The results indicate that most of users thought the video was of adequate quality and smooth. Question 13 asked users about their overall feeling of the UI layout of the player controls. The results indicate that positive opinions are balanced with negative opinions. In questions 14 to 21, users were asked whether or not the player features were easy to use, and on average approximately 75% of users (those chosen 4 and 5) thought the features were easy to use. **As a result, the overall usability has an above average score.**

Principle component analysis (PCA) was conducted to illustrate the distribution pattern of users with respect to the scores given to different types of questions. In this analysis, two principle component factors are used. The factor loadings (i.e. variable and factor correlation) are plotted in Figure 6-28 (see Appendix A for the original data). Questions 1, 2, 3, 5, 8, and 12 are closely related, plotted on the top right part of the figure. Questions 4, 6 and 13 to 20 are another closely linked group in the bottom right part of the figure. In general, the right part of the figure indicates relatively higher scores for these questions. For some other questions, e.g., questions 9 and 10, the right part of the diagram indicates relatively lower scores.

In Figure 6-29, each user's factor scores are plotted in a factor coordinating system so that users' distribution can be clearly presented. **Among all users, approximately half of them tend to give higher scores than others for questions concerning the overall experience, player features learnability and UI designs** (see Figure 6-29 and Figure 6-28).

A further analysis is based upon users grouped by age, gender and Internet TV use frequency.

In Figure 6-30, the squared rectangle labels the daily Internet TV users. It can be clearly seen that over half of them are allocated on the right hand side of the diagram (i.e. 8/13). **This suggests that more people tend to give higher scores in this half of the group than for remainder of users.**

In Figure 6-31, users are labelled with an age range. **The distribution of the largest age group (26 - 35 denoted by squared rectangle) indicates that half of them give higher scores than the other half.** This is, in effect, consistent with the general conclusion from Figure 6-29.

Finally, Figure 6-32 illustrates the user distribution by gender. Female viewers are labelled with a big squared rectangle. **This result suggests female viewers tend to give higher scores than male viewers.**

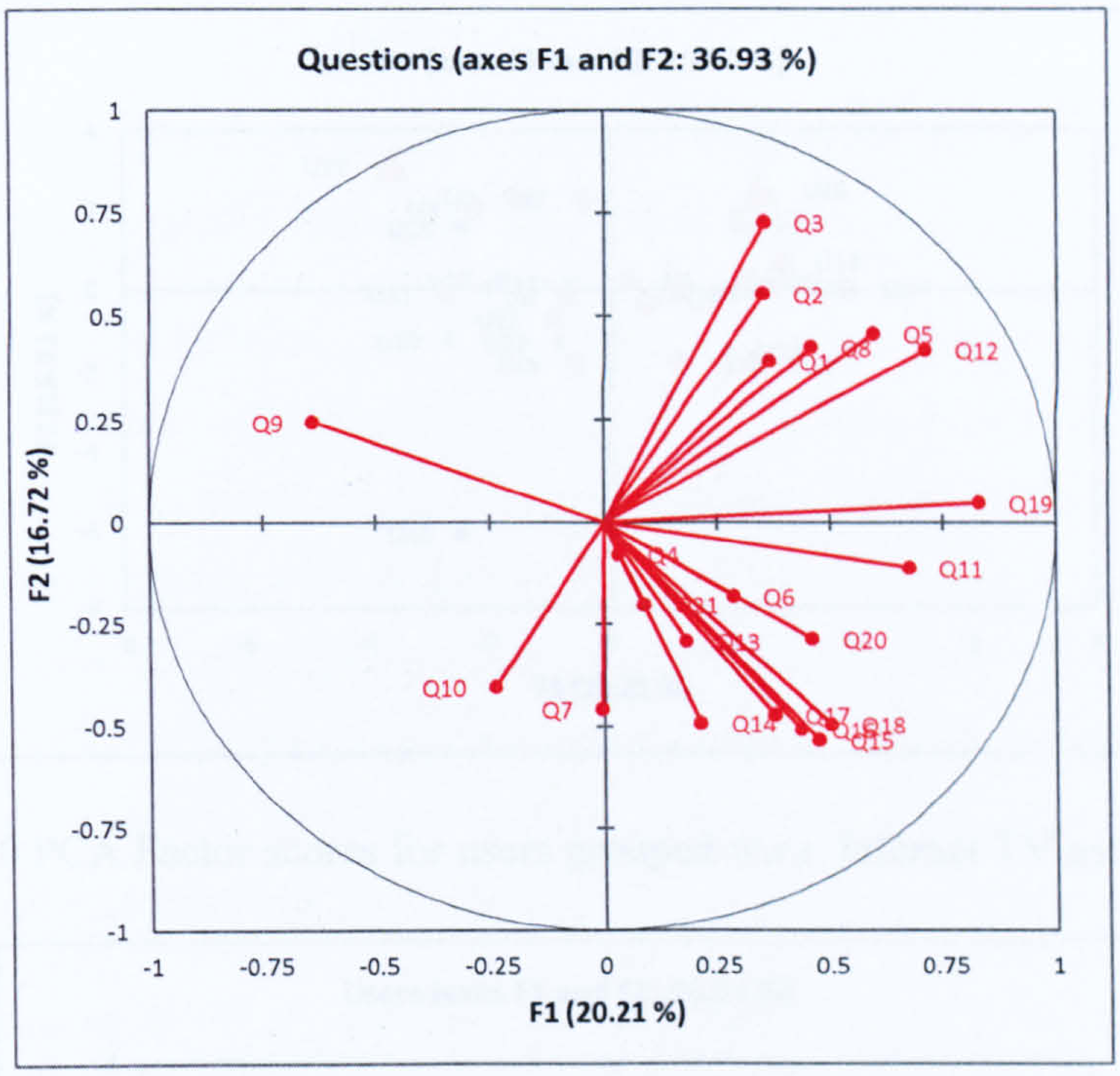


Figure 6-28 PCA factor loadings

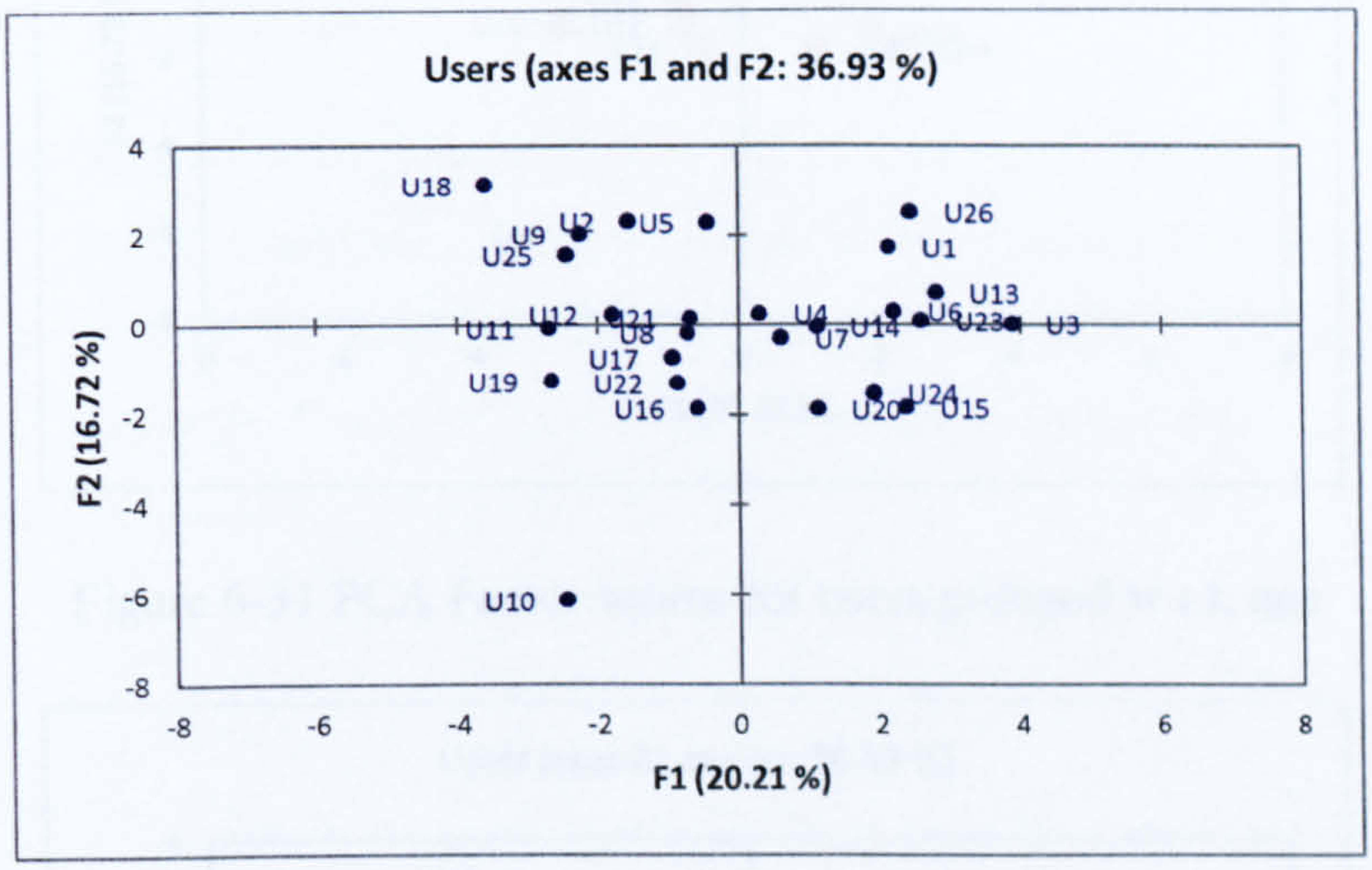


Figure 6-29 PCA Factor scores for each user

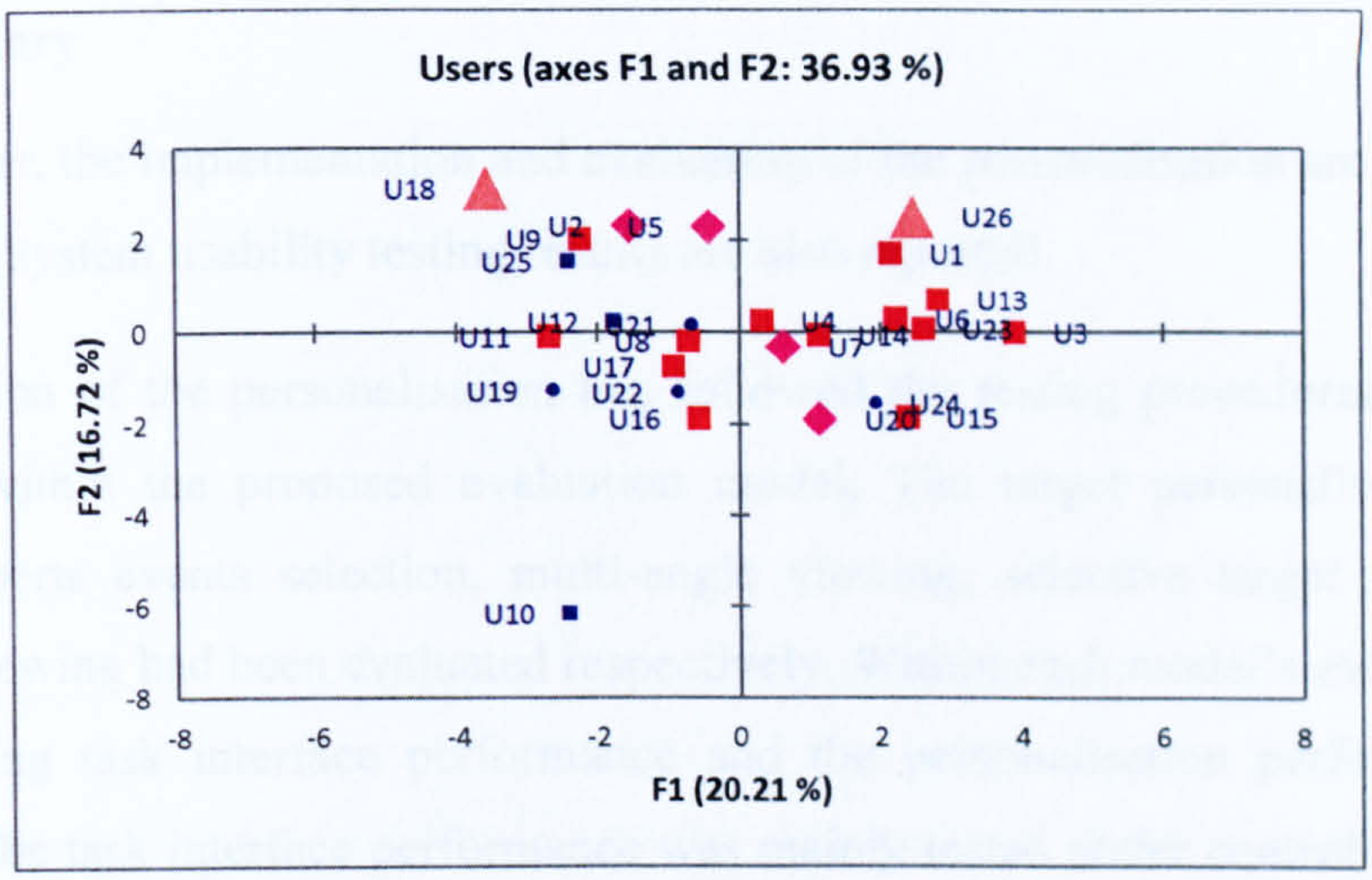


Figure 6-30 PCA Factor scores for users grouped w.r.t. Internet TV use frequency.

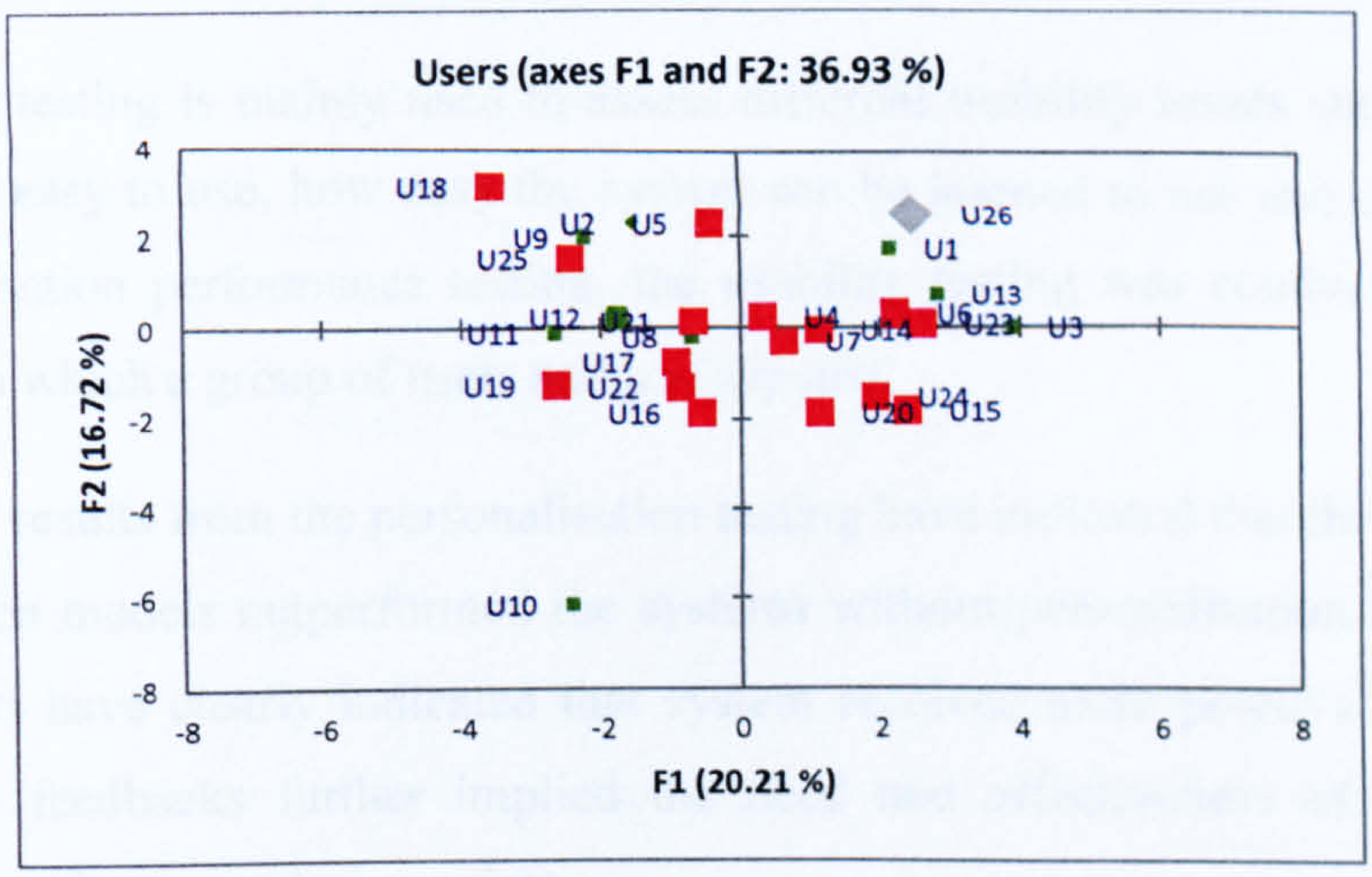


Figure 6-31 PCA Factor scores for users grouped w.r.t. age

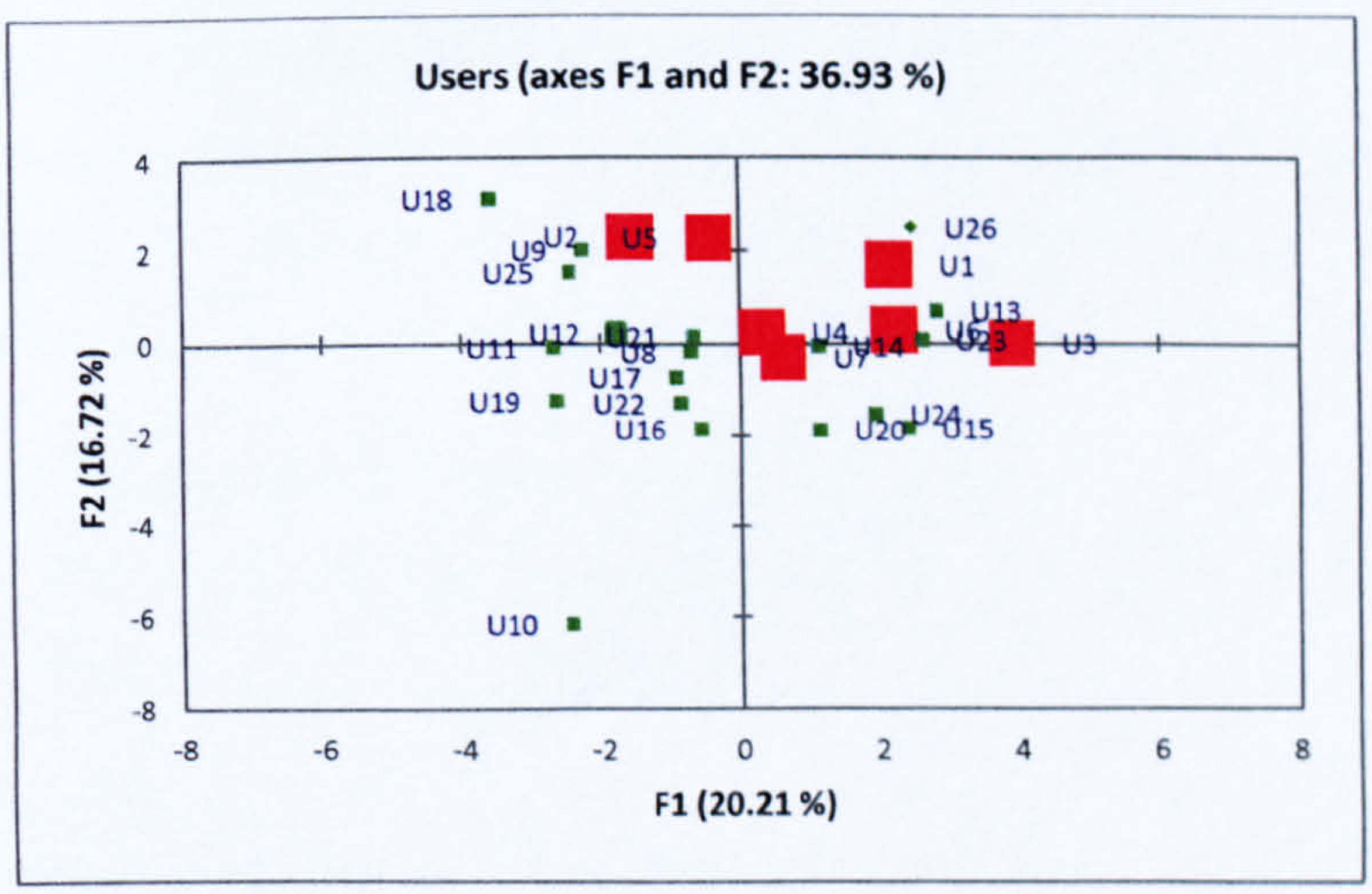


Figure 6-32 PCA Factor Scores for each users grouped w.r.t. gender

6.4 Summary

In this chapter, the implementation and evaluation of the personalisation are presented. In addition, the system usability testing results are also reported.

The evaluation of the personalisation has followed the testing procedures specified in Chapter 5 within the proposed evaluation model. The target personalisation models including sports events selection, multi-angle viewing, selective target zooming and time-shift viewing had been evaluated respectively. Within each model's evaluation, both the supporting task interface performance and the personalisation performance were evaluated. The task interface performance was mainly tested under controlled laboratory settings whereas the personalisation performance testing was conducted under real settings in which real users had participated.

The usability testing is mainly used to assess different usability issues such as whether the system is easy to use, how easy the system can be learned to use and etc. Similar to the personalisation performance testing, the usability testing was conducted under the real setting in which a group of users had participated.

The obtained results from the personalisation testing have indicated that the implemented personalisation models outperformed the systems without personalisation. The usability testing results have clearly indicated that system received more positive feedback and some of the feedbacks further implied the need and effectiveness of the proposed personalisation from users' viewpoints.

Chapter 7

7 Achievement and Future Work

This chapter summarises the chief research achievements of this work. At the end of this chapter, future work will be discussed.

7.1 Achievements

The achievements are threefold.

First, an advanced player offering advanced interactive features was researched and developed.

Second, a new user task model has been researched and developed. A personalisation model is advanced to enable a customised interactive A-V player to:

- model both individual and group users' preferences in the context of live sports events
- support active recommendations to groups with a low group re-clustering overhead
- model and leverage individual users' preferences to actively select targets to zoom in on, to use multi-angle views and to automatically adapt views of pre-determined sports incidents, through highlighting scenes and objects.

A system has been developed to actively adapt video content delivery to a network context. Bitrate adaptation handles bandwidth contention that occurs when two or more streams are concurrently displayed on a screen during zooming and during multi-camera viewing. Time-shift viewing can be interleaved into live events to support replays and objects highlights. In addition, a new evaluation model to assess personalisation performance with respect to consistency and scalability has been modelled and applied.

Third, a significant part of this author's work has contributed to key components of the My-e-Director 2012 project system[19]. A number of internal project deliverables and been published externally.

[19] See http://www.myedirector2012.eu for details.

7.1.1 Development Achievements

The novelty of this developed player can be highlighted from a comparison with other current A-V players.

Table 7-1 presents a comparison of the features for the custom A-V player reported in this thesis with the NBC's Sunday night football extra[20] and NBC Olympics Video Player[21]. Both the latter players are advanced industrial Web players that use Adaptive Streaming technology. Among the 14 target features for a next generation A-V player, the advanced A-V player in this thesis covers all of them while the other two industrial players only cover 50% of them.

Table 7-1 Player features comparisons

Player Features	This A-V Player	Sunday Night Football Extra	NBC Olympics Video Player
1. Multi-platforms (PC, Web, Win Mobile)	YES	YES	YES
2. Live content playback	YES	YES	YES
3. Adaptive Streaming	YES	YES	YES
4. Sports events selection	YES	NO	YES
5. Manual Zoom	YES	NO	NO
6. Manual Camera selection	YES	YES	NO
7. Manual Replay	YES	YES	YES
8. Manual Slow motion	YES	YES	YES
9. Installation requirements	NO	NO	NO
10. Personalised event selection	YES	NO	NO
11. Personalised Zoom	YES	NO	NO
12. Personalised Camera Selection	YES	NO	NO
13. Personalised Replay	YES	NO	NO
14. eDirector	YES	NO	NO

7.1.2 Personalisation Achievements

Personalisation of interactive tasks was structured, designed and modelled. The resulting personalisation mechanisms are implemented. They are also evaluated via a set of experiments including user tests. Table 7-2 presents the implemented personalisation mechanisms and their evaluative results.

[20] http://nbcsports.msnbc.com/id/26393211
[21] http://mediaevangelism.com/olympics/faq.aspx

Table 7-2 Personalisation implementation status and evaluation results

Personalisation	Brief Evaluative Results Summary
Personalised Event Selection	• Usability (from user feedback): • Easy to use and easy to learn • Personalisation: • Recommendation accuracy is better than system without personalisation • Recommendation accuracy is scalable in terms of both number of users and number of uses • Recommendation accuracy is in general consistent across users
Personalised Selective target zooming	• Usability (from user feedback): • Easy to use and easy to learn • Personalisation: • Zooming region of interest prediction precision is better than a system without personalisation • PP is scalable in terms of number of uses • PP is consistent across most of users
Personalised Multi-Angle Viewing	• Usability (from user feedback): • Easy to use and easy to learn • Personalisation: • Camera switching interval prediction precision is better than a system without personalisation • PP is scalable in terms of number of uses • PP is consistent across most users
Personalised Time-shift Viewing	• Personalisation (with modelled director expertise): • Expected incident highlighting precision and recall (the shorter the lead time, the higher the precision and recall) • Is able to highlight the critical object during a scene • Is able to replay the critical incidents (based upon 400m experiment)

7.1.3 Publications

The following lists the major publication achievement including My eDirector project deliverables and public publications.

Public Publications

- Wang, Z., Poslad, S. Personalised Live Sports Event Viewing on Mobile Devices. *The Third International Conference on Mobile Ubiquitous Computing, Systems, Services and Technologies (UBICOMM 2009)*, October 11-16, 2009 - Sliema, Malta

- Wang, Z., Poslad, S., Pearmain, A. A Collective Director for Highly Interactive Viewing of Live Sports Events. *EuroITV 2009 Workshop on Enhancing Social Communication and Belonging by Integrating TV Narrativity and Game-Play*, 3 June 2009, Workshop at EuroITV2009, Leuven, Belgium
- Patrikakis, C., Pnevmatikakis, A., Chippendale, P., Nunes, M., Santos Cruz, R., Poslad, S., Zhenchen, W., Papaoulakis, N., Papageorgiou, P., "Direct your personal coverage of large athletic events", *Multimedia, IEEE* , vol.PP, no.99, pp.1, 0

EU Project My-e-Director 2012 Internal Publications

- My-e-Director 2012 EU FP7 project deliverable D5.1 "User Terminal (Fixed, Mobile) and User Task Interface Definition Report", Available from http://www.myedirector2012.eu, submitted in 25th May 2009.
- My-e-Director 2012 EU FP7 project deliverable D2. 4 "System Architecture and Specifications", submitted on 31st July. 2009
- My-e-Director 2012 EU FP7 project deliverable D5.2 "Personal Preference Model Report", submitted on 25th Oct. 2009
- My-e-Director 2012 EU FP7 project deliverable D5.3 "Final report of the Application of the Personalisation Model", submitted on 31st Jan. 2011
- My-e-Director 2012 EU FP7 project deliverable D7.1 "Planning of Trials and Integration", submitted on 30th June. 2010
- My-e-Director 2012 EU FP7 project deliverable D7.2 "Report on Integration and Lab Trials", submitted on 30th October. 2010

7.2 Further Work

In this thesis, an advanced A-V player has been researched and developed that is, aimed at live sports events broadcast scenarios. The personalised interactions are not limited to athletic events but can apply to a broad range of sports events such as baseball, swimming etc. In addition, the user interfaces developed such as bitrate adaptation interface, video zoomable interface etc. can apply to an even wider range of video content. Although the A-V player is oriented towards Internet accessible PCs and laptop that use mouse-based input and PC screen output devices, it is also possible to use other types of the input or interaction such as 2D touch screens without changing the core parts of the task model. In order to extend and improve the work done in this thesis, personalisation can be further advanced in two ways as follows.

ICT and Physical world contexts may affect interactions and implicit user preferences

In this thesis, the ICT environment, i.e., video frame resolution and video stream bitrate, and the user model or profile, are considered as the main contexts that affect users' preferences to view live sports events. These can be extended. For example, those who view sports videos via mobile access devices require different types of interaction using 2D or 3D gestures rather than via pointer devices and they require simpler views for moving devices with smaller screens. In these cases, gesture patterns should be seen as implicit user preferences and they may vary across individual users. Obtaining such preferences requires the system to recognise, differentiate different patterns and relate these patterns to the A-V content. Other contexts such as the physical world context may also affect user interaction and user preferences and can be investigated. For example, users at live events with mobile devices, may wish to see event views that complement what they can directly see, when replaying event incidents at an event soon after they have occurred. To do this requires determining the position of users, in the face of an inability to use GPS in partially opened or closed sports stadiums, and the ability to model user views of an event and relate them to complementary multi-camera views that are available.

The design of rules used to adjust runtime online personalisation performance

From the experiments undertaken, it is clear that although the performance of personalisation is effective in most cases, it is not 100% in terms of critical metrics such as prediction precision. Nevertheless, the performance can be improved if suitable personalisation execution rules, i.e. rules about when to execute a personalisation mechanism, are given, as shown in this thesis. The design of such rules is challenging. Individual rules may not be sufficiently generic to personalise a range of user tasks and they may have a different impact on different users especially when personalisation is designed for individual users. The issue of personalisation execution rule design can be further investigated.

Appendix

Appendix A

Usability Descriptive Statistics

Questions	Users	Minimum	Maximum	Mean	Std. deviation
Q1	26	3.000	5.000	4.269	0.724
Q2	26	3.000	5.000	4.115	0.588
Q3	26	2.000	5.000	4.115	0.816
Q4	26	3.000	5.000	4.308	0.618
Q5	26	3.000	5.000	4.000	0.632
Q6	26	3.000	5.000	4.000	0.849
Q7	26	3.000	5.000	3.846	0.543
Q8	26	1.000	5.000	3.692	0.884
Q9	26	1.000	5.000	2.577	1.137
Q10	26	1.000	4.000	1.846	0.732
Q11	26	1.000	5.000	2.846	1.287
Q12	26	1.000	5.000	3.846	0.881
Q13	26	1.000	5.000	2.385	1.359
Q14	26	3.000	5.000	4.115	0.766
Q15	26	3.000	5.000	4.000	0.693
Q16	26	3.000	5.000	3.923	0.796
Q17	26	2.000	5.000	3.846	0.834
Q18	26	2.000	5.000	4.115	0.816
Q19	26	2.000	5.000	3.538	1.104
Q20	26	2.000	5.000	4.385	0.804
Q21	26	2.000	5.000	4.462	0.761

Appendix

Usability PCA Statistics

Factor loadings

Questions	F1	F2
Q1	0.364	0.392
Q2	0.351	0.554
Q3	0.354	0.728
Q4	0.033	-0.079
Q5	0.594	0.458
Q6	0.287	-0.181
Q7	-0.005	-0.458
Q8	0.457	0.426
Q9	-0.637	0.247
Q10	-0.238	-0.405
Q11	0.676	-0.112
Q12	0.708	0.419
Q13	0.182	-0.292
Q14	0.214	-0.493
Q15	0.475	-0.535
Q16	0.435	-0.509
Q17	0.377	-0.473
Q18	0.501	-0.496
Q19	0.830	0.049
Q20	0.460	-0.287
Q21	0.090	-0.202

Factor Scores

Users	F1	F2
U1	2.142	1.708
U2	-1.554	2.311
U3	3.885	0.004
U4	0.315	0.243
U5	-0.432	2.285
U6	2.213	0.293
U7	0.610	-0.313
U8	-0.718	-0.193
U9	-2.246	2.020
U10	-2.443	-6.142
U11	-2.689	-0.083
U12	-1.784	0.236
U13	2.810	0.708
U14	1.120	-0.057
U15	2.383	-1.842
U16	-0.576	-1.867
U17	-0.932	-0.747
U18	-3.576	3.132
U19	-2.649	-1.256
U20	1.129	-1.884
U21	-0.682	0.153
U22	-0.871	-1.305
U23	2.597	0.077
U24	1.927	-1.539
U25	-2.432	1.547
U26	2.453	2.508

Appendix B

For each question, a likert scale is used to qualify an answer where a score of '1' represents 'Not at all' and a '5' represents 'Very much').

- Question (Q) 1: How much did you enjoy using eDirector?

Results: 84% of the users enjoyed using the player (score 4 and 5).

- Question 2: Would you like to use eDirector again?

Results: 88% of users would like to use the player again (score 4 and 5).

- Question 3: Do you think that eDirector was easy to use?

Results: 81% of the users thought the player was easy to use (score 4 and 5), 15% of the users rated 3 while the rest of users rated it as 2.

- Question 4: Would you recommend eDirector to a friend?

Results: 92% of the users would be very likely to recommend the player to a friend (score 4 and 5).

- Question 5: Did you feel that the eDirector controls were simple to use?

Results: 19% of the users thought the controls were extremely simple (score 5) and 62% (score 4) of the users thought the controls were very simple, no user thought the controls were extremely difficult to use (i.e. score 1). Comment: this response is in-line with the question 3 response in which 81% of the users thought the whole system is easy to use.

- Question 6: Did you like being able to select your own camera angle to view?

Results: 66% of the users would like to select their own cameras (score 4 and 5). No user would not like to have camera switching option (i.e. score 1).

- Question 7: Did you like to have a choice of events to watch at any time?

Results: 75% of the users would like to have a choice of events (score 4 and 5). No one would like to have no choice.

- Question 8: Did you prefer to watch the footage chosen by professional directors?

Results: 61% of users prefer to watch the professional director's footage (score 4 and 5) whereas one user did not prefer it at all (score 1). Comment: A simple majority of viewers prefer to follow the footage from a human director but a sizable minority did not seem to. Supporting an option for both interactive (allowing users to direct their own footage) and automated view control (allowing the system to direct the footage) can complement the professional director's view.

- Question 9: I think that I would need help to be able to use eDirector

Results: 50% of the users think they would not need help (score 1 and 2). Comment: the results partly imply that the player seems new to many users but that it can be easily learned.

- Question 10: I needed to learn a lot of things before I could get going with eDirector

Results: 91% of the users thought they did not need to learn a lot of things before using the player (score 1 and 2).

- Question 11: I felt that the video was of high quality

Results: 35% of the users felt the video was of high quality (score 4 and 5) but 46% of the users felt the video was of low quality (score 1 and 2).

- Question 12: I felt that video playback was smooth

Results: 73 % of the users felt the video was smooth (score 4 and 5) and one user felt the video was not smooth (score 1 and 2).

- Question 13: I felt that the video controls got in my way

Results: 77% of the users felt the video control UI did not cause trouble (score 1 and 2)

- Question 14: How easy do you think the Basic playback controls (play, rewind, fast forward) of eDirector were to use?

Results: 77% of the users thought the basic playback controls were easy to use (score 4 and 5). No one thought it was not easy at all.

- Question 15: How easy do you think the advanced playback controls (live pause/play, rewind, fast forward) of eDirector were to use?

Results: 77% of the users thought the advance playback controls were easy to use (score 4 and 5). Comment; this agrees with the response to question 14.

- Question 16: How easy do you think the Event selection of eDirector was to use?

Results: 65% of the users thought the event selection were easy to use (score 4 and 5). No user thought it was not easy at all.

- Question 17: How easy do you think the Zoom of eDirector were to use?

Results: 63% of the users thought the zoom were easy to use (score 4 and 5). Two users thought it was less easy to use (score 2).

- Question 18: How easy do you think the Camera selection of eDirector was to use?

Results: 81% of the users thought the camera selection were easy to use (score 4 and 5). One user thought it was not easy at all.

- Question 19: How easy do you think the Event recommendations (recommendation to groups) of eDirector were to use?

Results: 54% of the users thought the event recommendation was easy to use (score 4 and 5). No user thought it was not easy at all.

Question 20: How easy do you think the Volume controls of eDirector were to use?

Results: 88% of the users thought the volume controls were easy to use (score 4 and 5). One user thought it was not easy at all.

- Question 21: How easy do you think the Full screen mode of eDirector were to use?

Results: 92% of the users thought the full screen mode were easy to use (score 4 and 5). One user thought it was not easy at all.

Appendix C

Demographics | Technology | **Sports Interests and Viewing Habits** | Quality/Price | Personalization

Which of the following sports/disciplines do you enjoy watching on television? (Select all that apply)

☐ Aquatics	☐ Fencing	☐ Shooting	☐ Archery	☐ Football
☐ Softball	☐ Athletics	☐ Gymnastics	☐ Table tennis	☐ Badminton
☐ Handball	☐ Taekwondo	☐ Baseball	☐ Hockey	☐ Tennis
☐ Basketball	☐ Judo	☐ Triathlon	☐ Boxing	☐ Modern pentathlon
☐ Volleyball	☐ Canoe	☐ Weightlifting	☐ Cycling	☐ Rowing
☐ Wrestling	☐ Equestrian	☐ Sailing		

Imagine you are at a sports event and have a mobile device with you. Which of the following options (on the mobile) would appeal to you?

☐ Ability to tag and track all athletes belong to a group, e.g., of the same nationality
☐ Ability to switch camera angles to view the event

☐ Ability to select and replay from a set of sports moments during the event
☐ Ability to watch replay in slow motion
☐ Ability to zoom in on user designated part of event, if quality of camera angle permits

☐ Ability to get summary information about a participant at any time
☐ Ability to get the event schedule at any time

☐ Audio commentary of the event
☐ Ability to replay or slow-motion of notable events
☐ Personalised Electronic Program Guide and Navigation to sports event

☐ Other, please specify: [ ]

Which of the following sports/disciplines would you like to watch live at the venue?

☐ Aquatics	☐ Fencing	☐ Shooting	☐ Archery	☐ Football
☐ Softball	☐ Athletics	☐ Gymnastics	☐ Table tennis	☐ Badminton
☐ Handball	☐ Taekwondo	☐ Baseball	☐ Hockey	☐ Tennis
☐ Basketball	☐ Judo	☐ Triathlon	☐ Boxing	☐ Modern pentathlon
☐ Volleyball	☐ Canoe	☐ Weightlifting	☐ Cycling	☐ Rowing
☐ Wrestling	☐ Equestrian	☐ Sailing		

Imagine you are at a sports event and have a mobile device with you. Which of the following options (on the mobile) would appeal to you?

☐ Ability to tag and track all athletes belong to a group, e.g., of the same nationality

☐ Ability to switch camera angles to view the event
☐ Ability to select and replay from a set of sports moments during the event
☐ Viewing the event from a camera angle that can't otherwise be seen by you, e.g., close-ups because you are situated too far from the event field

☐ Viewing athletes during the times they may otherwise be obscured from your view, e.g., they are on a part of the sports stadium that is temporarily hidden to you etc.
☐ No desire to use a mobile event at an event because

Imagine you are watching a long-distance athletics race, and could choose what view is presented. What type of view(s) would you like?

☐ Close-ups of specific athletes

☐ The front-runners
☐ The main pack

☐ An aerial view of the event
☐ I'd prefer just to view what the director has chosen

How would you like to control the user interface?

○ Use of arrow keys on keypad or keyboard

○ Use of mouse
○ Use of remote control

○ Use of touch pad
○ Use of touch screen

How would you prefer to learn how to use the interface?

○ A small set of simple controls but reduced interaction

○ Use of recorded demos to walk you through to show how the main feature work
○ Use of an on-screen help and hits triggered by indicating specific features

○ Access to electronic user guide off-line
○ Other, please specify: [ ]

Demographics | Technology | Sports | Media quality/price | Personalization

Imagine you could customise what you view (e.g. like in the example shown here [...]). To what extent would you make use of this facility on the following devices?

	Strongly Agree	Agree	Neither agree nor disagree	Disagree	Strongly disagree
Home TV set					
Desktop computer					
Laptop					
PDA, PocketPC, Palm					
Game Console					
Other					
Please, specify for other:					

For which of the following devices would you subscribe to a paid service that offers major sports events coverage in a personalised manner?

	Strongly Agree	Agree	Neither agree nor disagree	Disagree	Strongly disagree
Home TV set					
Desktop computer					
Laptop					
PDA, PocketPC, Palm					
Game Console					
Other					
Please, specify for other:					

For a content that costs you 2EUR, how much extra are you willing to pay for being able to watch the same content later (time-shift usage)?

- Not Interested
- 5% (10 cents) per view
- 10% (20 cents) per view
- 25% (50 cents) per view
- 50% (1 EUR) per view
- 75% (1,50 EUR) or more per view
- Other, please specify:

For a content that costs you 2EUR, how extra more are you willing to pay for being able to play the same content in other devices (device-shift usage)?

- Not Interested
- 5% (10 cents) per view
- 10% (20 cents) per view
- 25% (50 cents) per view
- 50% (1 EUR) per view
- 75% (1,50 EUR) or more per view
- Other, please specify:

Next

Bibliography

Aggarwal, Charu C., and Philip S. Yu. 2007. "On string classification in data streams", In *Proceedings of the 13th ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 36-45.

Akoumianakis, D., Grammenos, D. and Stephanidis, C., "User interface adaptation: Evaluation perspectives", *In User interfaces for All*, Stephanidis, C. (ed.). Erlbaum, pp. 339–352., 2001.

Ali, K., and Stam W., "TiVo: making show recommendations using a distributed collaborative filtering architecture", In *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 394-401, 2004.

Alpert, S. R. and Vergo, J. G., "User-centered evaluation of personalised web sites: What's unique?" *In Human Computer Interaction Research in Web Design and Evaluation*, Zaphiris, P. and Kurniawan, S. (eds). *Idea Group*, pp. 257–272, 2007.

Ambriola, V., Notkin, D., "Reasoning about interactive system", *Software Engineering, IEEE Transactions* , vol.14, No. 2, pp. 272-276, Feb 1988.

Annett, J. and Duncan K.D., "Task Analysis and Training Design", ERIC, 1967.

Ardissono, L., Alfred K., and Mark M. "Personalised Digital Television: Targeting Programs to Individual Viewers", *Human-Computer Interaction Series*, 2004.

Aroyo, L., Conconi, A., Dietze, S., Kaptein, A., Nixon, L., Nufer, C., Palmisano, D., Vignaroli, L., and Yankova, M. "NoTube – Making TV a Medium for Personalized Interaction", *EuroITV Leuven, Belgium*, 2009.

Assfalg, J. *et al.*, "Semantic annotation of soccer videos: Automatic highlights identification", *Comput. Vis. Image Understand.*, vol.92, pp. 285–305, 2003.

Baudisch, P., Good, N., Bellotti, V., and Schraedley, P., "Keeping things in context: A comparative evaluation of focus plus context screens, overviews, and zooming" , In *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI)*, pp. 259–266, 2002.

BBC internal Audience Research Reports from Olympic Games Athens (2004): *Audience Summary, Interest by Sport, Interactive TV Review*, from My-e-Director 2012 deliverable D2.1 "End-User Requirements", 2008 Available: http://www.myedirector2012.eu.

Beard, D. and Walker, J., "Navigational techniques to improve the display of large two dimensional spaces", *Behavior Info. Tech.* vol.9, No. 6, pp. 451–466, 1990.

Bederson, B. and Bolitman, A., "Does animation help users build mental maps of spatial information?", In *Proceedings of the IEEE Symposium on Information Visualization (InfoVis). IEEE Computer Society*, pp.28–35, 1999.

Bederson, B., Hollan, J., Perlin, K., Meryer, J., Bacon, D., and Furnas, G. 1996. "Pad++: A zoomable graphical sketchpad for exploring alternate interface physics", *J. Visual Lang. Comput.* vol.7, No. 1, pp. 3–31, 1996.

Bellekens, P., Houben, G. and Aroyo, L., "User Model Elicitation and Enrichment for Context-sensitive Personalisation in a Multiplatform TV Environment", *EuroITV'09*, 2009.

Benyon, D. and Murray, D.",Adaptive systems: From intelligent tutoring to autonomous agents", *Knowledge-Based Systems* vol.6 No. 4, pp. 197–219, 1993.

Benyon, D., Phil T., and Susan T. *Designing interactive systems: people, activities, contexts, technologies,* Harlow, England: Addison-Wesley. p.503 2005.

Bertini, M., Bimbo, A. D., and Torniai, C., "Multimedia enriched ontologies for video digital libraries", *International Journal of Parallel, Emergent and Distributed Systems*, vol.22, No. 6, pp. 407-416, 2007.

Bezdek, C. *Pattern Recognition With Fuzzy Objective Functions Algorithms*. New York: Plenum, 1981. ISBN:0306406713

Bondi, A. B. "Characteristics of scalability and their impact on performance", *In Proceedings of the 2nd international Workshop on Software and Performance (Ottawa, Ontario, Canada). WOSP '00. ACM*, New York, NY, pp. 195-203, 2000.

Boratto, L., Carta, S., Chessa, A., Agelli, M.,Clemente, M. L., "Group Recommendation with Automatic Identification of Users Communities", *Web Intelligence and Intelligent Agent Technologies, 2009. WI-IAT '09.* pp. 547-550, 2009.

Buchinger, S., Kriglstein, S. and Halvacs, H. "A comprehensive view on User Studies: Survey and Issues for Mobile TV", *EuroITV'09*, 2009.

Burghardt, C. and Kirste, T. "Inferring intentions in generic context-aware systems", In *Proceedings of the 6th international Conference on Mobile and Ubiquitous Multimedia. MUM '07, Vol.284. ACM, New York, NY*, pp. 50-54. 2007.

Campbell, D. and J. Stanley, *Experimental and Quasi-experimental Designs for Research*, Rand McNally college publishing company, Chicago, IL, 1966.

Card, S. K., Robertson, G. G., and Mackinlay, J. D. "The information visualizer, an information workspace", In *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI)*. ACM Press, pp. 181–186, 1991.

Cheung, G., Kanade, T., Bouguet, J.-Y., and Holler, M., "A real time system for robust 3d voxel reconstruction of human motions", *In CVPR*, pp. 714 – 720, 2000

Chin, D. N., "Empirical Evaluation of User Models and User-Adapted Systems", *User Modeling and User-Adapted Interaction Vol.*11, No.1-2, pp. 181-194, 2001.

Clarke, E. M., and Emerson, E. A., "Design and Synthesis of Synchronization Skeletons Using Branching-Time Temporal Logic." In *Logic of Programs, Workshop*, pp. 52-71, 1982.

Claypool, Mark, Phong Le, Makoto Waseda, and David Brown. "Implicit interest indicators", *In Intelligent User Interfaces*: pp. 33-40, 2001.

Cockburn, A., Karlson, A., and Bederson, B. B., "A review of overview+detail, zooming, and focus+context interfaces", *ACM Comput. Surv.* 41, 1 (Dec. 2008), pp. 1-31, 2008.

Darnell, M. J., "How do people really interact with TV?: naturalistic observations of digital tv and digital video recorder users", *Computers in Entertainment*, p. 10, 2007

De Ávila, P. M. and Zorzo, S. D., "A personalised TV guide system: an approach to interactive digital television", In *Proceedings of the 2009 IEEE international Conference on Systems, Man and Cybernetics* San Antonio, TX, USA, October pp. 11 - 14, 2009.

De Vel, O., and S. Nesbitt. "A Collaborative Filtering Agent System for Dynamic Virtual Communities on the Web", Working notes of Learning from. Carnegie Mellon University, 1998.

Dix, A., Janet F., Gregory D. A., and Russell B., *Human-computer interaction.* 1st ed. New York: Prentice Hall, 1993. ISBN: 0134372115.

Druin, A., Stewart, J., Proft, D., Bederson, B., and Hollan, J. 1997. "KidPad: A design collaboration between children, technologists, and educators". In *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI)*, pp. 463–470, 1997.

Dzida, W. and Freitag, R. Making Use of Scenarios for Validating Analysis and Design. *IEEE Transactions on Software Engineering.* Vol.24 No.12 December, pp.1182-1195, 1998.

Ekin, A., Tekalp, A. M., and Mehrotra, R., "Automatic soccer video analysis and summarization", *IEEE Trans. Image Processing*, Vol.12, No. 5, pp. 796–807, 2003.

Eronen, L., "Combining Quantitative and Qualitative Data in User Research on Digital Television", *In Proceeding of the 1st Panhellenic Conference PC HCI '01*, pp. 51–56, 2001.

Faltings B., Pu P., and Torrens M., "Designing Example-critiquing Interaction", *IUI 2004*, Madeira – Funchal, Portugal, pp. 22-29, 2004.

Fischer, P., Nurnberger, A., "Adaptive and multimodal interaction in the vehicle", *Systems, Man and Cybernetics, 2008. SMC 2008.* pp.1512-1516, 2008.

Foina, A.G., Badia, R.M., El-Deeb, A., Ramirez-Fernandez, F.J., "Player Tracker - a tool to analyze sport players using RFID", *Pervasive Computing and Communications Workshops* pp.772-775, 2010

Francois, D., Dautenhahn, K., Polani, D., "Using real-time recognition of human-robot interaction styles for creating adaptive robot behaviour in robot-assisted play", *Artificial Life, 2009. ALife '09.* pp.45-52, 2009.

Furnas, G. and Zhang, X., "MuSE: A multiscale editor". In *Proceedings of the ACM Symposium on User Interface Software and Technology (UIST)*, pp. 107–116, 1998.

Gallahue, David L., Frances Cleland Donnelly, and David L. Gallahue. 2003. *Developmental physical education for all children. Champaign*, IL: Human Kinetics. pp. 571–573. ISBN 0736033882.

Granic, A., Glavinic, V., "Adaptive systems and interaction: the design of personalised interaction in computer-based education", *Computational Cybernetics, 2005. ICCC 2005.* pp. 291- 296, 2005.

Green, T.R.G. and Benyon, D.R. The skull beneath the skin: entity-relationship modeling of information artefacts. *International Journal of Human-Computer Studies*, 44(6), pp. 801-828, 1996.

Guan Y., "Automatic Optimal View Selection for Natural HCI", *Image and Signal Processing, 2009. CISP '09.* pp.1-5, 17-19, 2009.

Gutwin, C., "Improving focus targeting in interactive fisheye views", In *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI).* ACM Press, pp. 267–274, 2002.

Hallberg, J., Svensson, S., Ostmark, A., Lindgren, P., Synnes, K., Delsing, J., "Enriched media-experience of sport events", *Mobile Computing Systems and Applications, 2004. WMCSA 2004. Sixth IEEE Workshop on* , pp. 2- 9, 2-3, 2004.

Herbrich, R., Graepel, T., and Campbell, C. 2001. Bayes point machines. *J. Mach. Learn. Res*, pp. 245-279, 2001.

Hesselman, C. and Shepherd, K. "Sharing Enriched Interactive TV Experiences with the iNEM4U Software Framework" (invited paper), *Workshop on Enhancing Social Communication and Belonging by Integrating TV Narrativity and Game-Play, co-located with the 7th European Interactive TV Conference (EuroITV2009),* Leuven, Belgium, 2009.

Hewett, T., *ACM SIGCHI curricula for human-computer interaction.* ACM. 1992., ISBN 0897914740.

Hölbling, G., Thalhammer, A., and Kosch, H. "Content-based tag generation to enable a tag-based collaborative tv-recommendation system", In *Proceedings of the 8th international interactive Conference on interactive TvandVideo.* EuroITV '10. ACM, New York, NY, pp. 273-282, 2010.

Holland, P. *The Television Handbook*, 2nd ed. London: Routledge, 2000, ISBN: 0415212812.

Höök, K., "Designing and evaluating intelligent user interfaces", *In Proceedings of the 4th international Conference on intelligent User interfaces*. IUI '99. ACM, New York, NY, pp. 5-6, 1999.

Höök, K., "Evaluating the utility and usability of an adaptive hypermedia system", *IUI '97*, Orlando, USA, 1997.

Höppner F., Klawonn, F. Kruse, R. and Runkler T., *Fuzzy Cluster Analysis*. West Sussex, U.K.: Wiley, 1999.

Hornbaek, K. and Hertzum, M., "Untangling the usability of fisheye menus", *ACM Trans. Comput. Hum.Interact.* vol.14, No. 2, p. 6, 2007.

Hornbaek, K., Bederson, B., and Plaisant, C., "Navigation patterns and usability of zoomable user interfaces with and without an overview", *ACM Trans. Comput.-Hum. Interact. Vol.9*, No. 4, pp. 362–389, 2002.

House, G., Sakamoto, S., Tajima, J., "Multiple object labeling in video sequences without skill", *Content-Based Access of Image and Video Libraries,* pp. 101-105, 1998

Huang Q., Hu J., Hu W., Wang T., Bai H., Zhang Y., "A Reliable Logo and Replay Detector for Sports Video", *Multimedia and Expo,* pp. 1695-1698, 2007.

Hudson, Scott E., Bonnie E. John, Keith Knudsen, and Michael D. Byrne."A tool for creating predictive performance models from user interface demonstrations". *In Proceedings of the 12th annual ACM symposium on User interface software and technology*, pp. 93-102, 1999.

Inamoto N., Saito H., "Virtual Viewpoint Replay for a Soccer Match by View Interpolation From Multiple Cameras", *Multimedia, IEEE Transactions*, vol.9, No. 6, pp. 1155-1166, 2007.

iNEM4U Consortium., 2009. Home. Available: http://www.inem4u.eu. Last accessed 6 Jan 2010.

Injae L., Myungseok K., Seyoon J., Kyuheon K., "Interactive Multi-View Visual Contents Authoring System", *Multimedia and Expo,2006 IEEE International Conference on*, pp. 189-192, 2006.

Jameson A. and Smyth B., "Recommendation to groups", *The Adaptive Web*, LNCS, pp. 596–627, 2007.

Jameson, A., Baldes, S., Kleinbauer, T., "Two methods for enhancing mutual awareness in a group recommender system", In: *Proceedings of the International Working Conference on Advanced Visual Interfaces*, Gallipoli, Italy pp. 447–449, 2004

Janez Z., Mladen S. "Real-time viewer feedback in the iTV production", In *Proceedings of the seventh european conference on European interactive television conference* (EuroITV '09). ACM, New York, NY, USA, pp. 149-152, 2009.

Jembere, E., M. O. Adigun, and S. S. Xulu. "Mining Context-based User Preferences for m-Services Applications", In *Proceedings of the IEEE/WIC/ACM International Conference on Web Intelligence*, pp. 757-763, 2007.

Joachims, Thorsten. "Optimizing search engines using clickthrough data", In *proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 133-142. 2002.

Jogalekar, P. P., and Woodside, C. M. "A Scalability Metric for Distributed Computing Applications in Telecommunications." In *Proceedings of the Fifteenth International Teletraffic Congress (ITC-15)* pp. 101-110, 1997.

John, B. "Information processing and skilled behaviour". In *Carroll, J.M. (ed.), HCI Models, Theories and Framework*, 2003

Johnson, N., Samuel K., Balakrishnan N., *Continuous Univariate Distributions* , Vol. 2 (Second Edition, Section 27). Wiley, 1995 ISBN 0471584940.

Jong D., M. D. T. and Schellens, P. J. "Reader-focused text evaluation. An overview of goals and methods", *Journal of Business and Technical Communication* 11(4), pp. 402–432, 1997.

Jumisko-Pyykko, S., Weitzel, M., and Strohmeier, D. "Designing for user experience: what to expect from mobile 3d tv and video?", *In Proceeding of the 1st international Conference on Designing interactive User Experiences For TV and Video,* ACM, New York, NY, pp. 183-192, 2008.

Kameda, Y., Koyama, T., Mukaigawa, Y., Yoshikawa, F., Ohta, Y., "Free viewpoint browsing of live soccer games", *Multimedia and Expo, 2004. ICME '04. 2004 IEEE International Conference*, pp. 747-750, 2004

Katsarakis, N., Pnevmatikakis, A., "Event detection in athletics for personalised sports content delivery", *Image Analysis for Multimedia Interactive Services, 2009. WIAMIS '09. 10th Workshop on*, pp. 280-283, 2009.

Kim, M., J. Lim, K. Kang, and J. Kim. "Agent-Based Intelligent Multimedia Broadcasting within MPEG-21 Multimedia Framework", *ETRI Journal*, pp. 136-148, 2004.

Klein, C. and Bederson, B. "Benefits of animated scrolling", In *Extended Abstracts of the ACM Conference on Human Factors in Computing Systems (CHI)*. ACM Press, pp. 1965–1968. 2005.

Knoche, H., Papaleo, M., Sasse, M. A., and Vanelli-Coralli, A., "The kindest cut: enhancing the user experience of mobile tv through adequate zooming", *In Proceedings of the 15th international Conference on Multimedia. MULTIMEDIA '07.* ACM, New York, NY, pp. 87-96, 2007.

Kosara, R., Miksch, S., and Hauser, H., "Semantic depth of field", In *Proceedings of the IEEE Computer Graphics and Applications, Symposium on Information Visualization (InfoVis)*, IEEE Computer Society Press, pp. 97–104, 2001.

Kripke, S. A., "Semantical considerations on modal logic", *Naming and necessity*, London: Oxford University Press, 1971.

Law, E. L., Roto, V., Hassenzahl, M., Vermeeren, A. P., and Kort, J. "Understanding, scoping and defining user experience: a survey approach", In *Proceedings of the 27th international Conference on Human Factors in Computing Systems CHI '09*. ACM, New York, NY, pp. 719-728, 2009.

Lawrence, R. D., G. S. Almasi, V. Kotlyar, M. S. Viveros, and S. Duri., "Personalisation of supermarket product recommendations", *The International Journal of Data Mining and Knowledge Discovery (Special Issue on Applications of Data Mining to Electronic Commerce)*, Vol.5, pp. 11-32, 2001.

Leahu, L., Phoebe S., and Michael M.,"Interactionist AI and the promise of ubicomp, or, how to put your box in the world without putting the world in your box", In *Proceedings of the 10th international conference on Ubiquitous computing*, pp. 134-143, 2008.

Lee S., Koo J., Chung Y., "Commercial Advance Technology with a DVDR System", *IEEE Transactions on Consumer Electronics*, vol.53, No. 2, pp. 461-466, 2007.

Leo M., D'Orazio T., Trivedi M., "A multi camera system for soccer player performance evaluation", *Distributed Smart Cameras, 2009. ICDSC 2009.* pp.1-8, 2009.

Leu J., Tsai C., Yi C., "Improving Adaptive Streaming Service across Wired/Wireless Networks", *Mobile Data Management: Systems, Services and Middleware, 2009. MDM '09. Tenth International Conference*, pp. 614-618, 2009

Li G., Franco, J.S., Pollefeys, M., "Multi-object shape estimation and tracking from silhouette cues", *Computer Vision and Pattern Recognition, 2008. CVPR 2008.* pp. 23-28, 2008.

Li L., Fengpei G., Qingwei Z., Yonghong Y., "Detecting cheering events in sports games", *Education Technology and Computer (ICETC), 2010 2nd International Conference*, pp. 223-227, 2010.

Li, L., Duan, L., Huang, Q., Du, J., and Gao, W., "A generic approach to classify sports video shots and its application in event detection", In *Proceedings of the First international Conference on internet Multimedia Computing and Service* ICIMCS '09. ACM, New York, NY, pp. 208-212., 2009.

Lim, K. Y. and Long, J. B. *The Muse Method for Usability Engineering.* Cambridge University Press, 2009, ISBN: 0521474949.

Lou, J., Cai, H., and Li, J., "A real-time interactive multi-view video system", In *Proceedings of the 13th Annual ACM international Conference on Multimedia.* ACM, New York, NY, pp. 161-170, 2005.

Ma, Y. and Zhang, H., "Motion pattern-based video classification and retrieval", *EURASIP J. Appl. Signal Process*, pp. 199-208, 2003.

Maltz, D., Ehrlich, E., "Pointing the way: Active collaborative filtering", *In CHI'95 Human Factors in Computing Systems,* 1995.

Mannila, H., Hannu T., and Verkamo A., "Discovery of frequent episodes in event sequences", *Data Mining and Knowledge Discovery* pp. 259--289, 1997.

Manzato, M., Coimbra, D.and Goularte, R, "Multimedia Content Personalisation Based on Peer-Level Annotation", *EuroITV'09*, 2009.

Martinez, A., Arias, J., Vilas, A., Garcia Duque, J., Lopez Nores, M., "What's on TV tonight? An efficient and effective personalised recommender system of TV programs", *IEEE Transactions on Consumer Electronics*, vol.55, No. 1, pp. 286-294, 2009.

Masthoff J. and Gatt. A. "In pursuit of satisfaction and the prevention of embarrassment: affective state in group recommender systems", *User Modelling and User-Adapted Interaction*, pp. 281–319, 2006.

Menon, N., Page, M., Watt M., and Bell, S., "Mobile Data Services: a Selection of Key Findings", In *Proceeding Of Mobinet '05*, 2005.

Middleton, S. E., Shadbolt, N. R., and Roure, D., "Ontological user profiling in recommender systems", *ACM Transaction. Information System*, pp. 54-88, 2004

Millerson, G. 1999. *Television Production*, Woburn, MA: Focal Press, ISBN: 0240514920.

Millerson, Gerald. 2001. *Video production handbook. Boston*, MA: Focal Press, ISBN: 174237056X.

Min X., Lakshmanan, L.V.S., Wood, P.T., "CompRec-Trip: A composite recommendation system for travel planning", *Data Engineering (ICDE), 2011 IEEE 27th International Conference on*, pp. 1352-1355, 2011

Miyauchi, S., Hirano, A., N. Babaguchi, and Kitahashi, T., "Collaborative multimedia analysis for detecting semantic events from broadcasted sports video", in *Proc. Int. Conf. Pattern Recognition*, vol.2, pp. 1009–1012. 2002.

MOBILE3DTV Consortium., 2009. Home. Available: http://sp.cs.tut.fi/mobile3dtv/. Last accessed 6 Jan 2010.

Montero, F., Víctor L., Jean V., and María L., "Solving the Mapping Problem in User Interface Design by Seamless Integration in IdealXML", In *Proc. of 12 th Int. Work-shop on Design, Specification, and Verification of Interactive Systems DSV-IS'2005* pp. 13-15, 2005.

Montgomery, D. C., and Runger, G. C., *Applied Statistics and Probability for Engineers*. John Wiley and Sons, 2010.

Morhee M., Tessens L., Luong H.Q., Prades-Nebot J., Pizurica A., Philips W., "A Distributed Coding-Based Content-Aware Multi-View Video System", *Distributed Smart Cameras, 2007. ICDSC '07.* pp. 355-362, 2007.

Mori, G, F P., and Santoro, C., "CTTE: support for developing and analyzing task models for interactive system design", *Software Engineering, IEEE Transactions* Vol.28, No.813, p.797, 2002.

Muntean G. M., Perry P., and Murphy L., "A new adaptive multimedia streaming system for all-IP multi-service networks", *IEEE Trans. Broadcasting*, vol.50, No. 1, pp. 1–10, March 2004.

My-e-Director 2012 Consortium, 2009., Home. Available: http://www.myedirector2012.eu, Last accessed 6 Jan 2010.

My-e-Director 2012 deliverable D5.1.1 "User Terminal and User Task Interface Definition Report", 2008, Available: http://www.myedirector2012.eu.

My-e-Director 2012 deliverable D7.3 "Report on Field Trials", 2011, Available: http://www.myedirector2012.eu.

Myers, Jerome L., and Well, A. *Research design and statistical analysis*. Mahwah, N.J.: Lawrence Erlbaum Associates, pp. 508, 2003. ISBN 080584037.

MYMEDIA Consortium., 2009. Home. Available: http://www.mymediaproject.org., Last accessed 6 Jan 2010.

Nakajima, T., Satoh, I., "Personal home server: a software infrastructure for supporting spontaneous and personalized interaction in home computing environments", *Consumer Communications and Networking Conference, 2005. CCNC. 2005 Second IEEE* , pp. 245- 250, 3-6 Jan, 2005

Nerkrasovski, D., Bodnar, A., Mcgrenere, J., Guimbertiere, F., and Munzner, T., "An evaluation of pan andzoom and rubber sheet navigation with and without an overview", In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2006.

Nichols, David M. "Implicit rating and filtering", *In Proceedings of the Fifth DELOS Workshop on Filtering and Collaborative Filtering*, pp. 31-36, 1997.

Norcio, A.F. and Stanley, J., "Adaptive human-computer interfaces: a literature survey and perspective", *Systems, Man and Cybernetics, IEEE Transactions*, vol.19, No. 2, pp. 399-408, 1989.

North, C. and Shneiderman, B., "Snap-together visualization: A user interface for coordinating visualizations via relational schemata", In *Proceedings of the Working Conference on Advanced Visual Interfaces (AVI 2000)*. ACM Press, pp. 128–135, 2000.

NoTube Consortium., 2009 Resources. Available: http://www.notube.tv/. Last accessed 6 Jan 2010.

Patterson, D. J., L. Liao, D. Fox, and H. Kautz. "Inferring High-Level Behavior from Low-Level Sensors". *Lecture Notes in Computer Science.* (2864): pp. 73-89. 2003.

Patterson, D. J., L. Liao, K. Gajos, M. Collier, N. Livic, K. Olson, S. Wang, D. Fox, and H. Kautz. "Opportunity Knocks: A System to Provide Cognitive Assistance with Transportation Services", *Lecture Notes in Computer Science.* (3205): pp. 433-450. 2004.

Perry, M., Juhlin, O., Esbjörnsson, M., and Engström, A., "Lean collaboration through video gestures: co-ordinating the production of live televised sport", In *Proceedings of the 27th international Conference on Human Factors in Computing Systems. CHI '09. ACM,* New York, NY, pp. 2279-2288, 2009.

Ping, S., Hong, Y., "An Item Based Collaborative Filtering Recommendation Algorithm Using Rough Set Prediction", *Artificial Intelligence, 2009. JCAI '09. International Joint Conference*, pp. 308-311, 2009.

Plumlee, M. D. and Ware, C, " Zooming versus multiple window interfaces: Cognitive costs of visual comparisons", ACM Trans. Comput.-Hum. Interact. Vol.13, No. 2, pp. 179-209, 2006.

Poslad S., *Ubiquitous Computing: Smart Devices, Environments and Interactions*. 1st ed. Wiley, p. 169, 2009.

Preece, Jenny, *Human-computer interaction*. Wokingham, England; Reading, Mass.: Addison-Wesley Pub. Co, 1995.

Prem M. and Vikas S., "Recommender Systems". In *Encyclopedia of Machine Learning*, Claude Sammut and Geoffrey Webb (Eds), Springer, 2010.

Recio-Garcia, J. A., Jimenez-Diaz, G., Sanchez-Ruiz, A. A., and Diaz-Agudo, B. "Personality aware recommendations to groups", In *Proceedings of the Third ACM Conference on Recommender Systems.* ACM, New York, NY, pp. 325-328, 2009

Ren, R. and Jose, J. M. "Attention guided football video content recommendation on mobile devices", *In Proceedings of the 2nd international Conference on Mobile Multimedia Communications*, MobiMedia '06, Vol.324. ACM, New York, NY, pp. 1-5, 2006.

Renaud, B., Eric, L., "Browsing Zoomable Treemaps: Structure-Aware Multi-Scale Navigation Techniques", *IEEE Transactions on Visualization and Computer Graphics*, pp. 1248-1253, 2007.

Rendle, S., Freudenthaler, C., Gantner, Z., Schmidt-Thieme, L., "BPR: Bayesian Personalized Ranking from Implicit Feedback",*In Proceedings of the 25th Conference on Uncertainty in Artificial Intelligence*, 2009.

Rice, M. and Alm, N., "Designing New Interfaces for Digital Interactive Television Usable by Older Adults", *Computers in Entertainment*, Vol.6, No. 1: pp. 1–20, 2008.

Rich, E., "Users are individuals: Individualizing user models", *International Journal of Man-Machine Studies,* Vol.18, No. 3, pp. 199-214, 1983.

Robertson, I. T., "Human information-processing strategies and style", *Behaviour and Inform. Tech.* Vol.4, No.1, pp. 19-29, 1985.

Rodgers, J. L. and Nicewander, A. W. "Thirteen ways to look at the correlation coefficient". *The American Statistician*, pp. 59-66, 1988.

Rodriguez, D., Goldmann, L., Ongkittikul, S., Karaman, M., Worrall, S., Sikora, T., Kondoz, A., "A system for personalised human computer interaction", *ELMAR, 2008.* pp. 439-442, 2008.

ROLE Consortium., 2009., Home. Available: http://www.role-project.eu., Last accessed 6 Jan 2010.

Sarah M., Eliza P., Nick T., Howard W., "Adapting Pervasive Environments through Machine Learning and Dynamic Personalisation", *International Symposium on Parallel and Distributed Processing with Applications*, pp. 395-402, 2008

Schalleck, R., Bober, M., and Drewes, H, "Design of an audience voting system for the Olympic games", *In CHI '04 Extended Abstracts on Human Factors in Computing Systems* (Vienna, Austria, April 24 - 29, 2004). CHI '04. ACM, NewYork, NY, pp. 1622-1625, 2004.

Schmidt, A., "Implicit Human Computer Interaction Through Context", *Personal Technologies*, pp. 191-199, 2000.

Seitz, S., Curless, B., Diebel, J., Scharstein, D., and Szeliski, R.. "A comparison and evaluation of multi-view stereo reconstruction algorithms", *In CVPR*, 2006.

Shelley B., Simone K., and Helmut H., "A comprehensive view on user studies: survey and open issues for mobile TV", In *Proceedings of the seventh european conference on European interactive television conference* ACM, New York, NY, USA, pp. 179-188, 2009.

Singh, G., Mohammed, I. A., and Sasi, S., "Real-time boundary detection for cricket game", In *Proceedings of the 3rd Australasian Conference on interactive Entertainment.* Australasian Conference on Interactive Entertainment, vol.207, Murdoch University, Murdoch University, Australia, pp. 23-27, 2006.

Soui, M., Benjaafar, I., Abed, M., Ghedira, K., "PES: Personalisation and evaluation system based on multi-agents approach: Application in transport information", *Intelligent Systems, 2008. IS '08. 4th International IEEE Conference*, vol.3, pp.23-2-23-6, 6-8, 2008

Sousa J., Poladian V., Garlan D., Schmerl B. and Shaw M., "Task-based Adaptation for Ubiquitous Computing", *IEEE Trans. on Systems, Man, and Cybernetics, Part C*, Vol.36 No. 3, pp. 328-340, 2006.

Spiliopoulou M., "Web usage mining for site evaluation: Making a site better fit its users", *Communications of ACM*, Vol.43, No. 8, pp. 127-134, 2000.

Srikant R. and Agrawal R., "Mining Generalized Association Rules", In *Proceedings of the 21th International Conference on Very Large Data Bases (VLDB '95)*, Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, pp. 407-419, 1995.

Stanton, N. A. "Human error identification in human-computer interaction", In *the Human-Computer interaction Handbook: Fundamentals, Evolving Technologies and Emerging Applications*, pp. 371-383, 2003

Strintzis, M. G. et al., "Knowledge representation for semantic multimedia content analysis and reasoning", In *Proceedings of the European Workshop on the Integration of Knowledge, Semantics and Digital Media Technology*, 2004.

Sugiyama, K., Hatano, K., and Yoshikawa, M. "Adaptive web search based on user profile constructed without any effort from users", In *Proceedings of the 13th international Conference on World Wide Web (New York, NY, USA, May 17 - 20, WWW '04*. ACM, New York, NY, pp. 675-684, *2004*.

Sung Y. J., Hong J., Kim T. "A formal model for user preference", *Data Mining, 2002. ICDM 2002. Proceedings,* pp. 235- 242, 2002

Svoen, B., "Consumers, Participants, and Creators:Young People's Diverse Use of Television and New Media", *Computers in Entertainment*, Vol.5, No. 2, 2007.

Tam, R. Chung-Man, David Maulsby, and Angel R. Puerta, "U-TEL: a tool for eliciting user task models from domain experts." *In Proceedings of the 3rd international conference on Intelligent user interfaces*, pp. 77-80. 1998.

Tan, Y. P., Saur, D. D., Kulkarni, S. R., and Ramadge, P. J, "Rapid estimation of camera motion from compressed video with applications to video annotation", *IEEE Trans. Circuits Syst. Video Technol.*, vol.10, pp. 133–146, 2000.

Tong Z., Vijay S Iyengar, and Pack Kaelbling."Recommender Systems Using Linear Classifiers", *Journal of Machine Learning Research*, pp. 313-334, 2002.

Trevisan, D. G., Monica G., Jean V., and Benoît M.,"Focus-based design of mixed reality systems." *In Proceedings of the 3rd annual conference on Task models and diagrams*, pp. 59-66, 2004.

Van Beusekom, M., Bignert, J., and Tasar, Ö., "SoMo: an automatic sound and motion sensitive audience voting system", In *CHI '04 Extended Abstracts on Human Factors in Computing Systems*, CHI '04. ACM, New York, NY, pp. 1680-1684, 2004.

Van Velsen, L., Van der geest, T., Klaassen, R., and Steehouder, M., "User-centered evaluation of adaptive and adaptable systems: A literature review", *Knowledge Engineering Review* pp. 261-281, 2008.

Vander Veer, G. C., Tauber, M. J., Waem, Y., and VanMuylwijk, B., "On the interaction between system and user characteristics", *Behaviour and Information Technology Vol.4 No. 4*, pp. 289-308, 1985.

Vuolle, M., Tiainen, M., Kallio, T., Vainio, T., Kulju, M., and Wigelius, H. "Developing a questionnaire for measuring mobile business service experience", In *Proceedings of the 10th international Conference on Human Computer interaction with Mobile Devices and Services. MobileHCI '08.* ACM, New York, NY, pp. 53-62, 2008.

Wages, R., Grunvogel, S.M., Zaletelj, J., Williams, C.M., Trogemann, G., "Future Live iTV Production: Challenges and Opportunities", *Automated Production of Cross Media Content for Multi-Channel Distribution, 2006. AXMEDIS '06. Second International Conference on*, pp. 325-328, 2006

Wang, J., Xu, C., Chng, E., Wah, K., and Tian, Q., "Automatic replay generation for soccer video broadcasting", In *Proceedings of the 12th Annual ACM international Conference on Multimedia.* ACM, New York, NY, pp. 32-39, 2004.

Wilcoxon, F., "Individual comparisons by ranking methods", *Biometrics*, pp. 80-83, 1945.

Winograd, Terry, "Shifting viewpoints: Artificial intelligence and human-computer interaction." *Artificial intelligence*. p. 1256, 2006.

Xie, L., Xu, P., Chang, S. F., Divakaran, A., and Sun, H., "Structure analysis of soccer video with domain knowledge and hidden markov models", *Pattern Recognit. Lett.*, Vol.24, p. 767, 2003.

Xiong Z., Radhakrishnan R., and Divakaran A., "Generation of sports highlights using motion activity in combination with a common audio feature extraction framework", *ICIP'03, vol.1*, pp. 5-8, 2003

Xu, C., Wang, J., Lu, H., Zhang, Y., "A Novel Framework for Semantic Annotation and Personalised Retrieval of Sports Video", *Multimedia, IEEE Transactions*, vol.10, No. 3, pp. 421-436, 2008

Xu. J., Zhang. L., Lu. H. and Li. Y., "The Development and Prospect of Personalised TV Program", *Proceedings of the IEEE 4th Internacional Symposium on Multimedia Software Engineering* (MSE'02), pp. 82- 89, 2002.

Yallow, E. "Individual differences in learning from verbal and figural materials", School of Educ. Stanford Univ., *Aptitudes Rcsearch Project Tech. Rep.* No. 13, 1980.

Yoshihama S., Chou P. and Wong D., "Managing Behaviour of Intelligent Environments", In *Proceedings of First IEEE International Conference on Pervasive Computing and Communications*. pp. 330-337, 2003.

Youngblood M., Holder L. and Cook D., "Managing Adaptive Versatile Environments", In *Proceedings of 3rd IEEE Int. Conf. on Pervasive Computing and Communications '05*, pp. 351-360, 2005.

Yu, Z., Zhou, X., Hao, Y., Gu, J., "TV program recommendation for multiple viewers based on user profile merging", *User Modelling and User-Adapted Interaction Vol.16, No. 1* pp. 63–82, 2006.

Zaletelj J., Savić, M. and Meža, M., "Real-time Viewer Feedback in the iTV Production", *EuroITV'09*, 2009.

Zhai, S., Conversy, S., Beaudouin-Lafon, M., and Guiard, Y., "Human on-line response to target expansion", In *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI)*. ACM Press, pp. 177–184, 2003.

Zhang, D. and Chang, S. F., "Event detection in baseball video using superimposed caption recognition", In *Proceedings of ACM Multimedia*, pp. 315–318, 2002.

Zhang, Y., Zhang, X., Xu, C., and Lu, H. "Personalised retrieval of sports video", In *Proceedings of the international Workshop on Workshop on Multimedia information Retrieval* (Augsburg, Bavaria, Germany, September 24 - 29, 2007), MIR '07. ACM, New York, NY, pp. 313-322, 2007.

Zhu, G., Huang, Q., Xu, C., Rui, Y., Jiang, S., Gao, W., and Yao, H. 2007. Trajectory based event tactics analysis in broadcast sports video. In *Proceedings of the 15th international Conference on Multimedia* (Augsburg, Germany, September 25 - 29, 2007). MULTIMEDIA '07. ACM, New York, NY, pp. 58-67.